# Correction of the classical *scid* mouse mutation by gene repair

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

AT THE UNIVERSITY OF LONDON

January 2013

BY

**Hayder Abdul-Razak**

CENTER FOR BIOMEDICAL SCIENCES

SCHOOL OF BIOLOGICAL SCIENCES

ROYAL HOLLOWAY UNIVERSITY OF LONDON

EGHAM, SURREY

## Declaration of Authorship

Unless otherwise stated in the text, all the work presented in this thesis was conducted by Hayder Abdul-Razak, working in the School of Biological Sciences, Royal Holloway University of London, between October 2008 and September 2012.

All the work is original unless acknowledged by references in the text. This work has not been submitted for any other degree at this or any other university.

Signed:

Date:

## Abstract

Gene addition strategies to correct inherited diseases are showing promise for treatment in the clinic, but an improved approach to treat both dominant and recessive genetic disorders would be to repair the mutant gene by homologous recombination-mediated gene targeting. Recently, gene targeting frequencies have dramatically increased through the development of (i) improved designer nucleases, including zinc finger nucleases (ZFN), able to introduce specific double-strand breaks at their target locus; and (ii) efficient DNA delivery tools, including integration-deficient lentiviral vectors (IDLVs). Even when using these state-of-the-art systems, gene repair frequencies in stem cells are moderate. To obtain proof-of-principle of phenotypic rescue in the haematopoietic system, a model in which corrected cells had selective advantage would be optimum. An *ex vivo* strategy to correct the classical *scid* mouse, a model of human DNA-dependent protein kinase catalytic subunit (DNA-PKcs, *PRKDC*) deficiency, is hereby presented. Donor templates to correct the *scid* point mutation and a ZFN were produced, optimised and incorporated into IDLVs. Specific ZFN activity in mouse *scid* fibroblasts and haematopoietic progenitors was demonstrated by *Cel*-I assay (which detects modifications introduced at the target site upon repair by non-homologous end-joining) and by deep sequencing. ZFN- and template-mediated gene repair of the *scid* mutation was demonstrated via the incorporation of a selection cassette and/or a diagnostic restriction site from the donor template into the targeted locus. In *scid* fibroblasts, gene repair led to DNA-PKcs activity rescue and increased resistance to DNA damage. In *scid* haematopoietic progenitors, gene correction was demonstrated only when the ZFN genes were delivered by integrating lentiviral vectors. Following an optimised *ex vivo* protocol, transplantation of potentially corrected *scid* haematopoietic progenitors into irradiated *scid* recipients has been carried out. Preliminary results of the *ex vivo* gene repair and transplantation experiment have indicated potential rescue of the T-cell compartment in a fraction of the *scid* transplant recipients. The results presented in this work highlight the potential of gene repair for future therapy in the haematopoietic system.

## Acknowledgment

In the name of ALLAH, Most Gracious, Most Merciful; praise be to ALLAH, the lord of the worlds and blessing be upon all his prophets and upon last prophets “Mohammad” blessings and peace be upon him.

Firstly, I would like to thank my supervisor Dr Rafael Yáñez for all his continuous support and for sharing his knowledge and expertise. I’d also like to thank my adviser Dr Jon Beauchamp for his helpful suggestions throughout the period of study. My thanks are also extended to Prof George Dickson for his kind support. I’d like to express my appreciation to Dr Chris Rider for his assistance as well.

A big thank to Prof Adrian Thrasher and the great team at Molecular Immunology Unit, ICH-UCL for their valuable practical assistance; in particular, Dr Steven Howe, Dr. Michael Blundell and Dr. Maria Alonso-Ferrero.

I’m very grateful to all previous and present members of Yáñez’s laboratory, it was nice accompany and good environment. I would like to especially thank my colleague Mrs Ngoc Lu-Nguyen for her personal support. Also I’m thankful to all members of Dickson’s laboratory for their continuous support. I’d particularly like to thank Dr Anita Le Heron, Dr Hanna Kymalainen and Dr.Martin Broadstock.

My everlastings thank and love goes to my family: Rana, Yaman, Yassar and Mayar who keep giving me all their love and support.

Finally, I would like to dedicate this thesis to my deceased beloved father who had always guided me to the right way; may Allah grant him eternal rest, Aameen.

## Abbreviations

| | |
|---|---|
| AAV | Adeno-associated virus |
| Ad | Adenoviral vectors |
| bp | Base pair(s) |
| BSA | Bovine serum albumin |
| CAR | Coxsackie and adenovirus receptor |
| CD | Cluster of differentiation |
| CFU | Colony forming unit |
| CHO cells | Chinese hamster ovary cells |
| CMV | Cytomegalovirus |
| CNS | Central nervous system |
| Cre | Recombinase enzyme (causes recombination) |
| DMEM | Dulbecco's modified Eagle's medium |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DNA-PK | DNA-dependent protein kinase |
| DNA-PKcs | DNA-dependent protein kinase catalytic subunit |
| ds | Double stranded |
| DSB | Double stranded breaks |
| *Env* | Envelope |
| EtBr | Ethidium Bromide |
| FCS | Foetal calf serum |
| FLP | Flippase recombination enzyme |
| Flt-3 | Fms-related tyrosine kinase 3 ligand |
| *Gag* | Group Specific antigen |
| GFAP | Glial fibrillary acidic protein |
| GFP | Green fluorescent protein |
| HEK-293T | Human embryonic kidney cell line, containing SV40 large T-antigen |
| HeLa | Cervical cancer cell line derived from Henrietta Lacks |
| HIV | Human immunodeficiency virus |
| HR | Homologous recombination |
| HSC | Haematopoietic stem cell |
| HSCT | Hematopoietic stem cell transplantation |
| IDLV | Integration-deficient lentiviral vector |
| InDel | Insertion/deletion |
| IPLV | Integration-proficient lentiviral vector |
| ITR | Inverted terminal repeat |
| Kb | Kilobase(s) |
| LAM-PCR | Linear amplification-mediated polymerase chain reaction |
| LTR | Long terminal repeat |
| MA | Matrix protein |
| MFI | Mean fluorescence intensity |

| | |
|---|---|
| MOI | Multiplicity of infection |
| m*Tert* | Mouse telomerase reverse transcriptase (immortalised fibroblasts) |
| MTT | Dimethylthiazol diphenyltetrazolium bromide |
| NHEJ | Non-homologous end joining |
| *Pac* | Puromycin resistance gene |
| PAGE | Polyacrylamide gel electrophoresis |
| PBMCs | Peripheral blood mononuclear cells |
| PBS | Phosphate buffered saline |
| pbs | Primer binding site |
| PCR | Polymerase chain reaction |
| PFA | Paraformaldehyde |
| PGK | Phosphoglycerate kinase |
| PIC | Pre-integration complex |
| PID | Primary immunodeficiency |
| Poly A | Polyadenylation |
| PPT | Polypurine tract |
| qRT-PCR | Quantitative real-time polymerase chain reaction |
| RBS | Rep-binding site |
| RCV | Replication-competent vectors |
| RE | Restriction enzyme |
| RNA | Ribonucleic acid |
| RRE | Rev responsive element |
| RSV | Rous Sarcoma virus |
| RT | Room temperature |
| S/MAR | Scaffold/Matrix attachment region |
| SCF | Stem cell factor |
| SCID | Severe combined immunodeficiency |
| SELEX | Systematic evolution of ligands by exponential enrichment |
| SFFV | spleen focus-forming virus |
| SIN | Self-inactivating |
| SSB | Single stranded breaks |
| SSC | Saline sodium citrate buffer |
| SV40 | Simian virus 40 |
| TALENs | Transcription activator-like effector nucleases |
| TAR | *tat* response element |
| TU | Transducing unit |
| V(D)J | Variable, Diverse, and Joining |
| VSV-g | Vesicular stomatitis virus glycoprotein G |
| WPRE | Woodchuck hepatitis virus post-transcriptional element |
| Wt | Wild-type |
| ZFN | Zinc finger nuclease |

# Table of Contents

# Chapter One

# Literature review

# 1 Literature review

## 1.1 Introduction to gene therapy

Gene therapy is an advanced form of molecular medicine that promises new treatments for the majority of inherited diseases (Verma and Weitzman, 2005). It can be defined as a deliberate alteration of the nucleic acid content of a patient's cells with aim to cure a disease or at least to arrest the development of a disease such as inherited and acquired diseases, in which a defective mutant gene is corrected via using a functional one (Hendrie and Russell, 2005; Verma and Weitzman, 2005). Gene therapy has the potential to revolutionise the approach of medical treatment; a single round of treatment might provide a lasting cure to what may otherwise need permanent treatment (Kohn, 2001). Thus, the main goal for gene therapy is to provide a lifelong treatment by either prevention of expression of unwanted genes or expression of a transgene, the gene of interest, at levels high enough to cure the disease or to at least to improve the patient clinically. As of August 2012, there were a total of 1843 gene therapy trails worldwide (http://www.wiley.co.uk/genetherapy/clinical, accessed on 16/08/2012 ). Currently, gene therapy is explored widely as a possible treatment for many serious disorders, like cancer, genetically inherited diseases, like haemophilia, cystic fibrosis, and severe combined immunodeficiency, and acquired diseases such as Parkinson's and Alzheimer's diseases.

A point of major concern related to the practical application of this approach is that of using an efficient delivery system capable of efficient gene transfer into a target cell without any problematic effects (Verma and Weitzman, 2005). In general, the target cells for the gene therapy are limited to two types; proliferative stem cells and highly differentiated long-lived cells. Undoubtedly the features of the targeted cells will significantly direct the choice of the delivery system. For example, it is important sometimes to ensure sustained expression of the transgene, a process that can be achieved by using delivery vectors that can integrate in the genome efficiently. While in other cases and especially when non-dividing cells are targeted, it is preferable to choose non-integrated vectors (Vink *et al*., 2009). In this regard, the introduction of genetic material into targeted cells is based on two concepts: the first is *ex vivo* gene introduction which involves taking away of the desired cells outside the body to achieve

*in vitro* cellular modification, by transduction for example, followed by re-administration of the modified cells back into the body. The second is *in vivo* gene introduction which involves direct introduction, local or systemic, of genetic material into target cells and tissues of the body (Mulligan, 1993; Wolff and Lederberg, 1994)

Two main strategies have been developed to achieve the ambitious aims of gene therapy. First "gene addition" involves complementing a faulty gene with a therapeutic one, using vectors that typically contain a functional expression cassette allowing persistence via an extrachromosomal episome or integration at non-homologous chromosomal sites (Hendrie and Russell, 2005); though the faulty gene remains intact (Nakayama, 2010). The second approach known as gene repair, site-specific repair, or gene targeting can be defined as a correction of the defective gene by introduction of specific genetic change using homologous vector sequence (Delenda, 2004). The first gene repair study was carried out in 1989 to correct mutated HPRT in murine animal model (Thompson *et al*., 1989).

Gene repair approach offers many advantages over gene addition, including:

- the defect in the target locus of interest can be corrected more precisely by gene repair.
- the proper copy number of the gene is preserved.
- gene repair, unlike gene addition, reduces the possibility of creating risky insertional mutations and oncogene activation which may be caused by random integration.
- gene repair requires the delivery of only the nonfunctional sequence of the faulty gene; therefore it is not related to the full size of the faulty gene whereas in gene addition the whole functional region needs to be delivered which occasionally will be inapplicable for large size genes.
- the corrected gene will be expressed under the control of its own normal regulatory endogenous chromosomal elements, resulting in precise correction of the defective gene that will repair the genetic disorder lastingly by re-establishing the normal function for the gene; leading to reduced possibility of non-physiological gene expression and long-term gene inactivation problems which normally accompany gene addition and,

- the ability of gene repair to correct both dominant and recessive mutations, whilst gene addition is limited to recessive defects only (Ellis *et al*., 2012; Levine *et al*., 2006; Yanez and Porter, 1998).

To date two main delivery systems have been used in gene therapy studies: viral and non-viral vectors. The non-viral vectors include for instance: direct microinjection, gene gun, liposome formulations, and synthetic peptides. In general, viral vectors represent the most successful gene delivery systems (Moldt *et al*., 2008). RNA viruses, retroviruses and lentiviruses, are exploited for integrative gene manipulations, whereas DNA viruses provided non integrative manipulations (Verma and Weitzman, 2005). Increasingly lentiviral vectors have become very useful tools in genetic manipulation, because of their unique ability to insert genetic cargo in both dividing and non-dividing cells (Moldt *et al*., 2008). Despite many successful applications that use viral vectors as efficient gene delivery systems, there are a number of drawbacks accompanying this approach. These issues include: maintaining and regulating of transgene expression (Cornu and Cathomen, 2007; Vigna and Naldini, 2000), random integration which may lead to insertional mutagenesis (Modlich *et al*., 2009; Philpott and Thrasher, 2007)), and relatively limited genetic packaging capacity (Philippe *et al*., 2006).

## 1.2 Viral Vectors

Viral vectors are either RNA or DNA viruses, they could be found as both integrating into the host DNA and non-integrating as an episomal element. Their use in gene therapy is based on harnessing their ability to deliver genetic material into transduced cells, thus some viral pathogens could be recruited as delivery vectors by removing their elements causing disease and employing their components needed for infection and replication. The nature of viral life cycle generally determines the protocols of production. As an example, adenovirus and adeno-associated virus (AAV) vectors are harvested by lysis of infected cells to release the viral particles because they are produced from transfections. One of the important features that typical viral vectors should possess is to be produced in a highly concentrated form to ensure stable and sustained gene expression. Additionally, viral vectors should be able to target the most suitable cell type for the disease. And most importantly, produced viral vectors should

have no pathogenic or unfavourable effects relevant to their transduction and finally have no or lower immunogenicity (Tomanin and Scarpa, 2004; Warnock *et al*., 2011).

The murine leukemia virus (MLV) from family retroviridae was the first virus to be used in gene therapy applications as viral delivery systems (Fischer and Cavazzana-Calvo, 2005). Interestingly they have remained useful in gene therapy because of their long-term expression since the genes they transduce are integrated into the genome of the target cells. Conversely, their use is restricted in *in vivo* gene transfer of non-proliferating cells, such as hepatocytes, myofibers, and neurons because of their nature to transduce proliferating cells only (Naldini *et al*., 1996). Currently, the introduction of lentiviral vectors into the gene therapy applications has limited the use of MLVs as the former vectors can transduce both proliferating and non proliferating cells.

### 1.2.1 Adenoviral vectors

Adenoviruses (Ads) are nonenveloped, icosahedron -shaped protein capsid (70-100 nm in diameter) that surrounded inner nucleic acid containing core (Raus *et al*., 2011). They were first isolated from human adipose tissue in 1953 (Warnock *et al*., 2011) and since then, Ads were isolated from other species. Ads are usually associated with mild respiratory symptoms, keratoconjunctivitis and gasteroenteritis (Khare *et al*., 2011). These viruses differ from retroviruses and lentiviruses in that they do not integrate into the host genome (Nakayama, 2010). This will lead to reduced risk of insertional mutagenesis; yet their use in gene therapy is restricted to non dividing cells as cell division will result in loss of the transgene (Stephen *et al*., 2010). Ads can efficiently transduce their target cells *in vivo* and are the most commonly used vectors in gene therapy (Shirakawa, 2008). The Ads-based vectors have been used in many gene modification trials like: retinal diseases and ocular malignancies such as retinoblastoma (Ildefonso *et al*., 2012); peritoneal fibrosis (Li *et al*., 2012b); cystic fibrosis (Hida *et al*., 2011); and cardiovascular diseases (Kawase *et al*., 2011).

Ad genomes are linear, double stranded DNA of about 36-40 kb; and as vectors they possess the ability to incorporate large therapeutic transgenes. Their DNA is organised into nucleosome-like structure and known to have inverted terminal repeats (ITR) on the DNA two ends (Figure 1.1) (Khare *et al*., 2011). These viruses are naturally transducing

their target cells via binding to cellular receptor “Coxsackie and adenovirus receptor” (CAR) which is mainly expressed by epithelial cells. The CAR receptors mainly mediate viral interaction with most of target cell types. So far, there are about 57 different human Ad serotypes have been isolated, which are divided into subgroups A to G. (Wong *et al*., 2012). However the serotypes 2 and 5 represent the most widely used Ads (Douglas, 2004).

Ad genes are organized into two main transcription regions: the early and late, and minor delayed region (Iva2) based on the time course of their expression before and after DNA viral replication, respectively (Warnock *et al*., 2011). Ads, like other DNA viruses, need to deliver their genome into nucleus to trigger viral replication. As a defence mechanism, cells can recognise spontaneous replication or foreign nucleic acids initiating DNA damage response and leading to cell cycle arrest and/or apoptosis. Therefore, DNA viruses have developed ‘early’ viral genes to block cellular antiviral machinery (Schreiner *et al*., 2012). These early genes are: E1A which necessary to stimulate early phase transcription enhancing the S phase of the target cell; E1B which required to prevent apoptosis allowing viral replication; E2 encodes viral proteins required for replication; E3 encodes proteins that inhibit cellular responses to viral infection; and finally E4 encodes different proteins required for viral DNA transcription (Warnock *et al*., 2011).

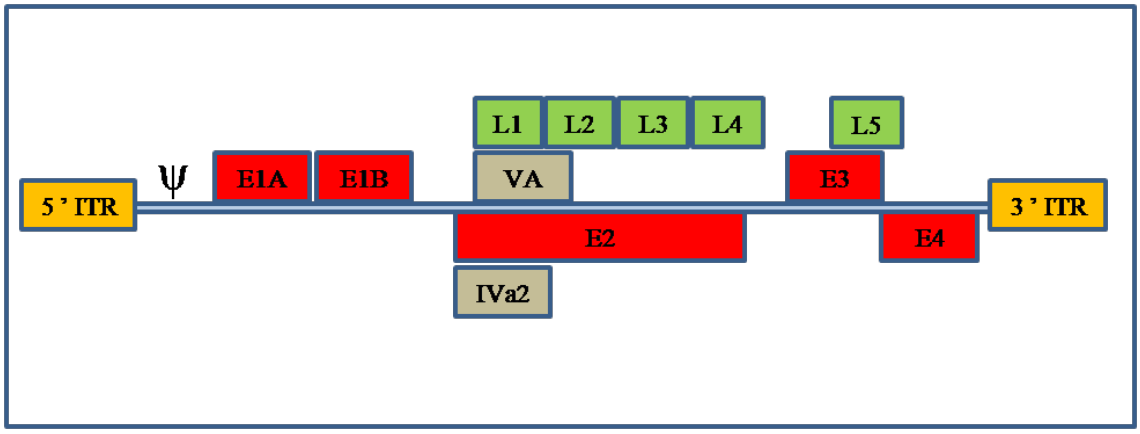



**Figure 1.1: Schematic representation of adenovirus genome organisation.** The inverted terminal repeats (ITR) are located on the DNA ends and in between them are the early genes (E1A, E1B, E2, E3 and E4), delayed gene (IVa2) and late genes (L1, L2, L3, L4 and L5). Ψ: RNA packaging signal. Diagram adapted from Verma and Weitzman, 2005.

The initial interaction between Ads and their host cells is facilitated by a high affinity attachment of the viral fiber knob domain with a CAR cell receptor. This is followed by beginning of endocytic uptake and viral internalisation via clathrin-mediated endocytosis through the penton base interaction with so-called alpha V integrins. Next, an early endosome containing virion will be formed and the viral compartments begin to dissociate in the low pH endosome environment. Then the viral particles will be released out of the endosomes into the cytoplasm via endosomal membrane endosomolysis by the action of viral specific proteins. The viral DNA will be moved by microtubules reaching the nuclear pore complex. Finally, the viral DNA will enter the nucleus by transportation through the nuclear pore (Nemerow *et al*., 2009).

The genetic modification of adenoviral genome and capsid has led to safe recombinant Ad-based vectors (Shirakawa, 2008). The Ad vectors have many advantages over other viral vectors such as: the relative ease of production and high functional vector titers achievable; their ability to infect non-dividing cells with broad target cell tropism; delivering of their double-stranded DNA genome into the nucleus in a form that is naturally episomal at high efficiency; and absence of mutagenic effect. Nevertheless, the most potential advantage of adenovirus vectors is their ability to package large capacity for DNA insertion (Hendrie and Russell, 2005; Li *et al*., 2012b; Verma and Weitzman, 2005). In contrast, their ability to stimulate strong immune and inflammatory reactions, and their promiscuity related to transduction of other different cells in addition to targeted cells, has limited their use (Kawase *et al*., 2011).

Ads are known to have lytic life cycle inside their host cells; therefore Ad-based vectors were developed by removing those viral genes that cause this undesirable effect. In this regard, Ad first generation vectors designed to be replication-deficient by the deletion of El early genes, which is essential for expression of E2 genes and other genes necessary for DNA synthesis and replication. The combined deletion of E1 and some of E3 genes that are not necessary for viral replication *in vitro* and involved in stimulation of viral immunogenicity, has led to development of Ad second generation vectors. These two deletions in E1 and E3 regions provide about 8 kb in the Ad vector genome to be used for insertion of DNA of interest, ensuring sustained expression of transgenes in Ad vector transduced-mammalian cells. Additional subsequent deletion of some E4 genes has introduced Ad third generation (Campos and Barry, 2007; Vorburger and Hunt,

2002). On the other hand, gutted (also called gutless or helper-dependent) adenoviral vectors are reported to be considerably safer compared to early generation adenoviral vectors. Unlike first- and second generation adenoviral vectors which still include some of the adenoviral coding sequences, gutted adenoviral vectors are devoid of all viral coding sequences. They can induce sustain transgene expression and possess significantly less inflammatory potential than early-generation Ad vectors (Ehrhardt and Kay, 2005; Toietta *et al*., 2005).

Human embryonic kidney (HEK) 293T cells were previously transduced with modified Ad5 genomic DNA establishing HEK 293T cell line which stably expressing Ad E1 genes. Hence E1-deficient Ad vectors can be produced and amplified in these 293T cell lines via trans-complementation (Graham *et al*., 1977). Currently Ads vectors are produced in 293T cell line at efficient titres, and purified using CsCl density gradient centrifugation.

### 1.2.2 Retroviral vectors

Retroviruses are enveloped viruses with ~ 100 nm in diameter; their genomes consist of two identical copies of RNA of varied sizes ~ 8-11 kb surrounded by a conical-shaped core (Goff, 2001). These viruses are considered the only truly diploid viruses and they are the only RNA viruses whose genome is produced without any participation of viral encoded polymerase depending only on cellular transcriptional machinery (Coffin *et al*., 1997). Retroviruses have the ability to efficiently integrate their genetic material into host chromosomes (Hendrie and Russell, 2005). They are known for their ability to promote reversal of genetic transcription from RNA to DNA using reverse transcriptase upon to transduction of host cell genome. Some retroviruses can cause various fatal human diseases, including acquired immunodeficiency syndrome (AIDS), inflammatory diseases and multiple types of cancer.

These viruses are subdivided into two categories, simple and complex, according to the organisation of their genome. Simple retrovirus carries only essential major genes, *gag*, *pol*, and *env*; while complex retrovirus encodes additional regulatory genes that play important roles in their replication and pathogenesis. In fact, all retroviruses possess the basic genome structure of centrally located main genes (*gag*, *pol*, and *env*) flanked by

long terminal repeats (LTR) at either ends of genome. The basic organisation of retroviral genome is showed in Figure 1.2. The role of LTRs is important for regulating transcription of the viral genes and integrating of viral DNA and initiating its synthesis (Shida, 2012). Retroviruses are further subdivided into seven genera; five are the alpha- to epsilon-retroviruses that are characterised by oncogenic abilities, and two other genera, the spumaviruses and lentiviruses (Lim, 2012). In general simple retroviruses, like gammaretroviruses and alpharetroviruses, transduce non dividing cells at reduced efficiencies; though lentiviruses show approximately no difference between dividing and non-dividing cells. So, some retroviruses are almost completely dependent on cell cycle progression, others show an intermediate dependence, and lentiviruses have nearly no dependence on cell proliferation (Yamashita and Emerman, 2006).

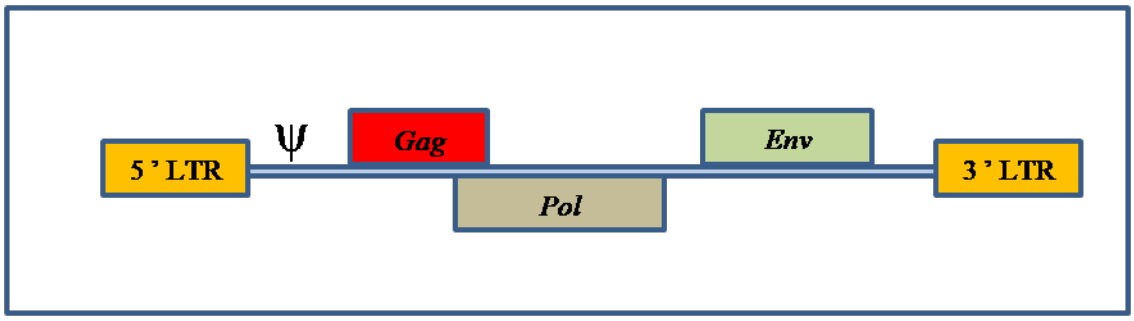



**Figure 1.2: Schematic representation of retroviral genome organisation.** The long terminal repeats (LTR) located on the DNA either ends and in between them structural genes (*gag*, *pol* and *env*), Ψ: RNA packaging signal. Diagram adapted from Delenda, 2004.

The *gag* region encodes genes for the matrix, capsid proteins, and nucleocapsid which are produced by proteolytic cleavage of the *gag* precursor protein; the *pol* region encodes genes for reverse transcriptase, protease, and integrase viral enzymes which are generated from the *gag-pol* precursor, and the *env* region encodes genes for the envelope glycoproteins and proteins needed for receptor recognition, which facilitate virus entry. The life cycle of retroviruses starts with binding of viral surface glycoproteins incorporated in the viral envelope to its receptor on the surface of the host cell. Any modifications in these glycoproteins will affect viral tropism, hence they mediate the viral tropism and they were exploited for vector pseudotyping. Then the entry of viral capsid containing the RNA genome into the cell will be mediated by endocytosis. Subsequently, the retroviral RNA genome will be transcribed into a double-stranded proviral DNA by the viral enzyme reverse transcriptase. The reverse

transcribed proviral DNA, reverse transcriptase, integrase, nucleocapsid, and other viral proteins will be combined together to form a preintegration complex (PIC). Then the PIC will be translocated into the nucleus, and viral DNA enters the nucleus as a part of the PIC. The integrase enzymes catalyse the integration of the provirus into the host chromatin. Next, the cellular machinery of transcription will be initiated by viral LTR and the proviral DNA will be transcribed into mRNA and genomic retroviral RNA. The viral mRNA then will be translated into glycoproteins and nucleocapsid proteins and viral compartments will be formed at the plasma membrane. Finally, the cell membrane of host cell buds out and new progeny virions will be released (Sundquist and Krausslich, 2012; Verma and Weitzman, 2005; Vile, 1992).

Retroviral-based vectors are designed for safety reasons by removal of those genes essential for viral replication and their genomes were additionally modified to encode therapeutic transgenes between the two LTRs. Retroviral replication-defective vectors were mainly recruited to create genetic modifications in stem cells for treatment of various human diseases, including blood disorders, diabetes, and neurological disorders (Lim, 2012). However, retroviral related insertional mutagenesis has emerged as potential side effects after retroviral gene therapy. In this regard, the previous assumption of random retroviral integration has recently changed; it is believed now retroviral integration does not occur randomly. Retroviruses posses the tendency to integrate close to transcription start sites of active genes inducing activation of a proto-oncogene or disruption of a regulatory gene (Fischer and Cavazzana-Calvo, 2005). As a result when retroviral integration occurs near oncogenes, normal cells can be transformed into cancerous cells, as demonstrated by the incidence of leukaemia in earlier gene therapy trials (Fischer and Cavazzana-Calvo, 2005; Fischer *et al*., 2012; Qasim *et al*., 2009).

### 1.2.3 Lentiviral vectors

Lentiviruses (LVs) are complex retroviruses, known with their abilities to transduce both dividing and non-dividing cells and maintain long-term transgene expression (Lesch *et al*., 2011). LV genomes, like other retroviruses, comprise the structural three genes of *gag*, *pol*, and *env*; however they additionally code for a set of accessory genes that required to ensure efficient viral transduction/ replication in the host cells. In

Human Immunodeficiency Virus Type -1 (HIV-1), the common member of lentiviruses, these genes includes two regulatory genes vital for viral replication: *rev* encodes regulator viral protein and *tat* encodes transactivator of transcription, and four accessory genes essential for *in vivo* rep viral replication and pathogenesis and not critical for virus growth *in vitro*; *vif* encodes viral infectivity factor, *vpr* encodes viral protein R, *vpu* encodes viral protein U and *nef* encodes negative factor (Matsuda *et al*., 2009). Additionally, LVs possess nuclear localization signals (NLS) which facilitate entry of PIC into nuclei of non dividing cells enabling LVs to transduce non dividing cells like neurons (Li *et al*., 2010). A schematic representation of the HIV-1 structure is illustrated in Figure 1.3 and diagrammatical representation of the HIV-1 genome organisation is shown in Figure 1.4 A.

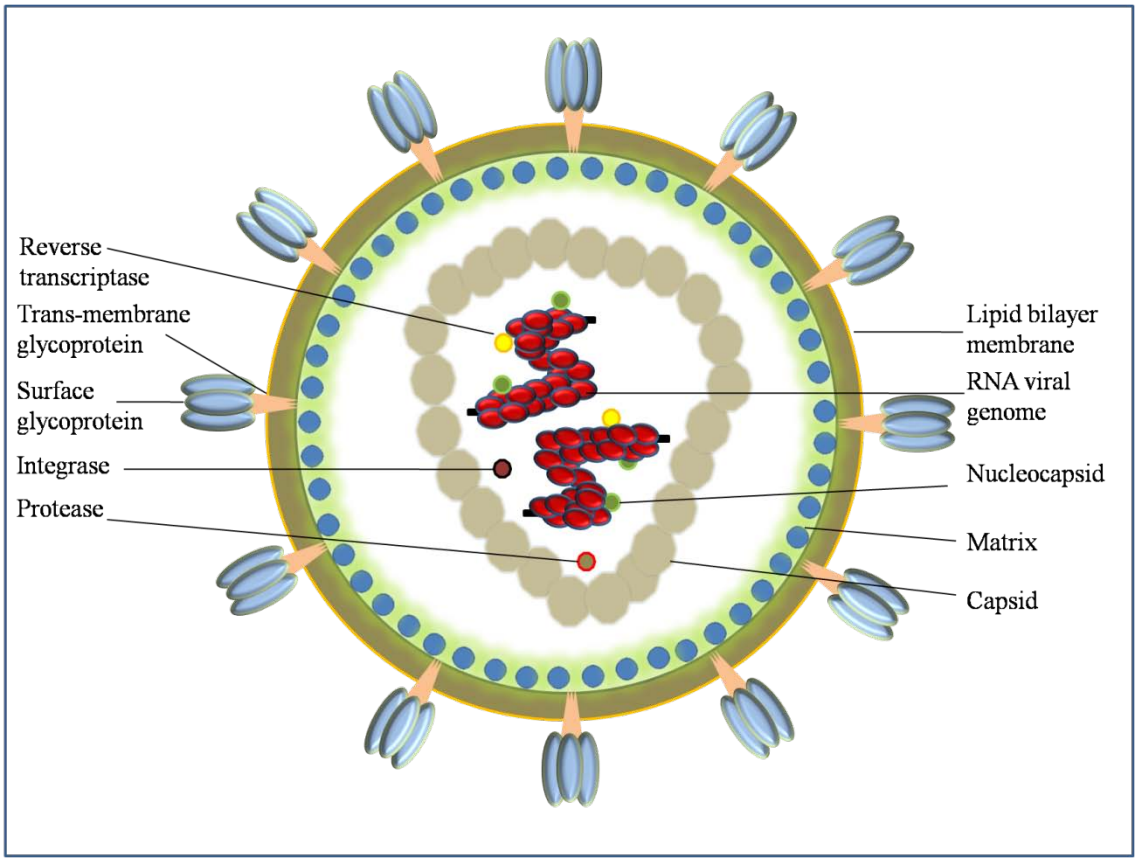



**Figure 1.3: Schematic representation of the HIV-1 structure.** General diagram of wild type HIV-1 structure showing a viral conical-shaped core including matrix, capsid and nucleocapsid surrounded by viral bilayer lipid envelope. Viral core also comprises viral essential enzymes: reverse transcriptase, integrase and protease. The envelope contains surface and trans-membrane glycoproteins essential for attachment entry to host cell. Adapted from Buchschacher and Wong-Staal, 2000.

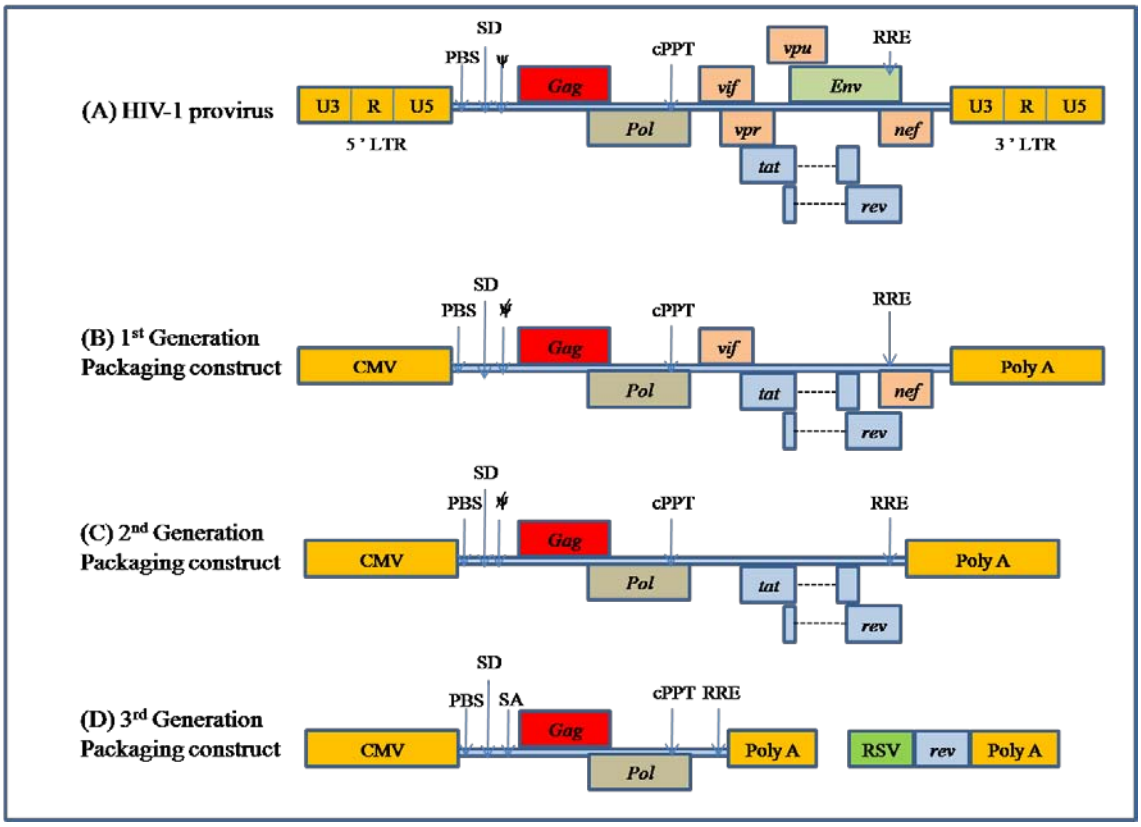



**Figure 1.4: Diagrammatical representation of developing of HIV-1 based lentiviral vector packaging constructs.** Panel A, gene organisation of typical HIV-1- provirus, panel B, C and D are first-, second- and third-generation types of packaging constructs comprise derived features of the HIV-1, respectively. The cytomegalovirus (CMV) promoter and polyadenylation signal (polyA) replaced viral 5' and 3' LTRs, respectively. The envelope was deleted and some of accessory proteins were mutated in first generation packaging construct keeping structural and most of accessory genes intact. Yet, all the accessory genes have been removed in second generation packaging construct. While subsequent modification in third generation packaging construct has led to remove *Rev* and *Tat* genes and to express the *Rev* gene under Rous Sarcoma Virus (RSV) promoter and a polyA from separated and non-overlapping construct. SA: splice acceptor, SD: splice donor site. The figure is not to scale and is adapted from Vigna and Naldini 2000, and Zufferey *et al*. 1998.

The lentiviral LTRs of the viral RNA comprise two sequence elements: the 5'LTR includes the repeated R and unique 5' (U5) region whereas the 3'LTR contains the unique 3' (U3) region in addition to a second copy of the 5'R region. During the reverse

transcription, portions of both LTRs will be duplicated and the reverse transcribed DNA will possess U3, R and U5 sequences at either ends. Upon integration of reverse transcribed DNA into host cell genome, the proviral DNA will act as a template for host transcription machinery during replication and formation of new viral particles (Buchschacher and Wong-Staal, 2000).

The basic structure of HIV-1 genome additionally includes c*is*-acting nucleic acid sequence. These non- translated elements are essential for viral packaging and reverse transcription and are mainly located in the highly structured region of 5' end. They contain: the *gag* leader sequence region located between 5' LTR and *gag* region which comprise the encapsidation signal (ψ), the main splice donor (SD) site and primer binding site (PBS) which is needed to initiate the reverse transcription process; central polypurine tract (PPT) located in the central region of the genome and in the 3' region which is required for DNA synthesis; and Rev-responsive element (RRE) which is located near the 3' end and is essential for RNA transport (Buchschacher and Wong-Staal, 2000).

#### 1.2.3.1 Development of lentiviral vectors for gene therapy

Lentivectors are LV based vectors modified for gene therapy applications by harnessing LVs natural abilities in transduction of both dividing and non-dividing cells. Continuous vector development has always focused on the production of efficient and safe vectors relying on minimising the chances of replication-competent vectors (RCVs) arising, insertional mutagenesis and oncogenesis, and ensuring at the same time efficient viral expression. The aim is to design a viral modified genome encoding a therapeutic gene cassette to replace the viral original genome. Accordingly, this development produces a non-pathogenic, non-replicative and abortive viral transduction, which introduces transgenes into the target cells and ultimately generates efficient expression from the recombinant vectors.

The main strategy used for developing lentiviral vectors was to eliminate genes responsible for pathogenesis. For example, in third generation lentiviral vectors this strategy has result in splitting the viral genome to be expressed from separated plasmids leading to separate structural and regulatory genes. In that case the viral structural genes

encode essential enzymes and proteins are expressed from *trans* acting elements in the packaging plasmids, whereas regulatory genes encode viral gene expression and packaging of vector genome are expressed from *cis* acting elements in the transfer plasmids. Then these two systems (*trans / cis*) are brought together by co-introduction into the producer cells. The principle of this approach aimed at increasing the recombination events needed for production of (RCVs) and eventually reduces the possibility of generating of virulent viruses (Kay *et al*., 2001).

In this connection, first generation lentiviral vectors were generated by separation of the packaging, envelope and vector cassettes and express them in *trans* from 3 separated plasmids. The *Env* was removed, some of accessory genes of packaging plasmid were mutated and most of other HIV-1 provirus features were maintained (Figure 1.4 B), keeping minimal *cis*-acting elements essential for packaging. The lentiviral vectors were pseudotyped with vesicular stomatitis virus glycoprotein G (VSV-G) which is carried by separated plasmid (Naldini *et al*., 1996). This approach intended to reduce the opportunities of RCVs generation by homologous recombination and the same time lentiviral vectors were produced inside producer cells upon cotransfection.

Along with *trans / cis* system, another strategy was applied to improve viral biosafety based on the deletion of non-essential viral genes leading to design of the second generation lentiviral vectors. The accessory genes *vif*, *vpr*, *vpu*, and *nef* encode important proteins to HIV-1 infection but not essential for viral replication and packaging. However, *rev* was considered vital for gene expression, not only in wild type viruses but in LV based vectors as well, while *tat* was considered essential only for those vectors based on transfer plasmids with HIV LTR, like pHR' family. Importantly, Zufferey *et al* showed that deletions of HIV accessory genes *vif ,vpr*, *vpu* and *nef* from the packaging plasmid, for safety reasons, did not result in reduction of vector production (Zufferey *et al*., 1997). Hence, of the nine genes in the original HIV pathogen, only four – *pol*, *gag*, *rev* and *tat* were considered crucial for vector production (Figure 1.4 C).

A further lentiviral system that became known as the third generation lentiviral vectors (Figure 1.4 D) was early developed by removing *tat* expression from the packaging construct. Additionally, the *gag*/*pol* and rev genes were segregated and expressed from two distinct constructs. Moreover, as from first generation lentiviral vectors the CMV,

as a constitutive promoter, replaced the viral 5' LTR, while the polyA signal replaced viral 3' LTR (Dull *et al*., 1998).

Additional improvements of lentiviral vectors safety have exploited the viral replication mechanism to generate self-inactivating (SIN) vectors. The HIV RNA contains promoter / enhancer sequences at LTR of both the 5' and 3' ends. During the reverse transcription the tRNA primer binds to the PBS and LTR start to extend firstly at the 5' end of the viral RNA, then transferring to the 3' end allowing reverse transcriptase to complete transcription in the 3' to 5' direction. Hence, this process 'copies' the 5' end of the LTR to the 3' end of the growing DNA strand and vice versa. This process has been exploited by introducing a deletion via removal of 400 base pairs from the U3 region comprising the TATA box in the 3' LTR, which after reverse transcription in the host target cell is flipped into the U3 region of the 5'LTR of the proviral DNA. As a result this deletion essentially eliminates the promoter function of the LTR and consequently prevents the expression of full-length vector RNA. Thereby the generated lentiviral vectors will be only capable of producing a single copy of double stranded DNA for each RNA molecule resulting in further improving of vector safety aspects. The lentiviral vectors produced by this strategy are called SIN vectors as their transgene expression will be available from only the internal promoter in the transfer construct and have even weaken abilities to produce RCVs. The significance of the deletion in the enhancer region of the 3' U3 of the LTR gene has made SIN vectors safer because it results in a transcriptionally inactive vector that can not be converted into a full length RNA and reduces likelihood of RCV regeneration. Eventually this will lead to reduce risk of tumorigenesis via promoter insertion (Logan *et al*., 2004; Miyoshi *et al*., 1998; Qasim *et al*., 2009; Zufferey *et al*., 1998).

The woodchuck hepatitis virus post-transcriptional regulatory elements (WPREs) are found to enhance transgene expression and transduction efficiency of lentiviral vectors. The placement of WPREs near the 3' untranslated region downstream of the transgene has been assumed to improve mRNA nuclear export and genomic transcription (Zufferey *et al*., 1999). This element act at post-transcriptional level, most possible to enhance the efficiency of polyadenylation of newly transcribed RNA leading to increase it is amount in the cells (Vigna and Naldini, 2000). It has been reported that

incorporation of WPREs in transgene construct could result in overall transgene expression by more than fivefold (Werner *et al*., 2004).

Pseudotyping of lentiviral based vectors with a variety of envelope proteins, commonly the VSV-G glycoprotein, has led to: widening viral tropism considerably to infect different species both *in vitro* and *in vivo*; enhancement of viral production techniques by stabilisation of vectors particles from shearing forces during ultracentrifugation leading to achieve high titres; and, in LVs particularly, facilitating the viral entry into host cells by endocytic pathway reducing the viral dependency on accessory proteins. Moreover, retroviral based vectors incorporation of combined synthetic and tissue-specific promoters has evoked sustained and regulate gene expression in vector host cells (Cockrell and Kafri, 2007; Mikkola *et al*., 2000).

The LV natural features together with improvements mentioned above have resulted in making the lentivectors attractive tools in gene therapy because:

- Their relative large packaging capacity (up to 8 Kb), which is a key feature in packaging large expression cassettes of tissue-specific promoters and transgenes,
- Their ability to transduce effectively both dividing an non-dividing cells,
- Their ability to integrate genes into targeted chromosomal DNA and subsequently lead to stable and sustained gene expression, and
- Their reduced immunogenicity compared to other vectors, making them possible considerable systemic delivery routes and reducing the likelihood of anti-vector antibody formation (Verma and Weitzman, 2005; Wanisch and Yanez-Munoz, 2009).

Lentiviral vectors are commonly produced by transient cotransfection of HEK 293T cells using recombinant constructs encoding transgenes, packaging elements, rev gene and sequences encoding Env glycoproteins (Figure 1.5), and harvesting lentiviral-containing supernatant. The Env VSV-G is used widely to replace HIV-1 envelope protein, providing highly stable capsid vectors; although a wide range of alternative glycoproteins have the ability to attach to lentivectors's membrane. It has been suggested that VSV-G attach to phosphatidylserine on host cell membrane (Burns *et al*., 1993; Kutner *et al*., 2009).

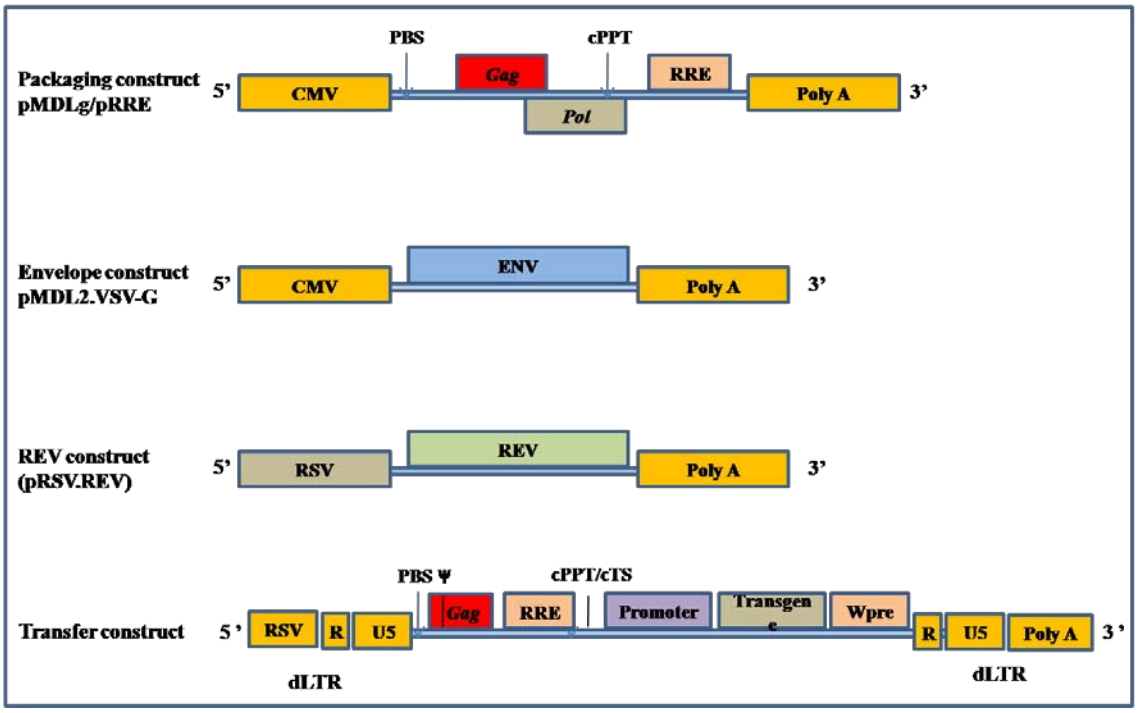



**Figure 1.5: Diagrammatical representation of recombinant constructs used for lentiviral vector production.** Different constructs used to produce lentiviral vectors based on third generation system. The figure is not to scale.

### 1.2.3.2 Integration-deficient lentiviral vectors

Integration of viral DNA copy into the host cell genome is an essential step of the life cycle of retroviruses as it is ensures long-term expression of the retroviral genes allowing a productive viral propagation. Retroviral integration has been confirmed to be non-random and at virus-specific locations (Ciuffi, 2008). As an example, gamma-retroviral vectors are known by their ability to integrate close to the LMO2 proto-oncogene, resulting in triggering and then overexpression of this gene leading eventually to oncogenesis. Lentiviruses have the tendency to integrate their genomes near or within transcriptionally active host cell genes, having, unfortunately, a risk of insertional mutagenesis and subsequent oncogenesis. In general lentiviruses have strongly attenuated, but not abrogated, oncogenic potential (Montini *et al.*, 2006). From safety point of view, this could cause a major problem for gene therapy approach as reported earlier (Qasim *et al.*, 2009). In order to circumvent possible related insertional mutagenesis and to reduce its risk, the integrase-deficient lentiviral vectors (IDLVs) are introduced. The development of IDLVs is considered a further improvement to boost

HIV-1 derived vectors biosafety, as they still keep the earlier characteristics of integrating proficient lentiviral vectors (IPLV) (Philippe *et al*., 2006).

The viral enzyme integrase plays a critical role in viral integration, reverse transcription and nuclear import. Upon virus entry into the host cell, integrase will combine with other elements to form PIC which will mediate nuclear entry, and eventually the integrase will mediate proviral DNA integration into the chromosomal DNA genome (Arhel *et al*., 2007). Integrase catalyses the hydrolysis of two terminal nucleotides from each end of HIV-1 DNA leaving a 5' dinucleotide overhang and 3' hydroxyl exposed group, this process called "3 processing". Consequently, the integrase employ ed the exposed 3' hydroxyl group to cut opposite strands of chromosomal DNA facilitating the connection of the viral exposed DNA 3 ends to the generated 5′ overhangs via phosphodiester bond. In the end, the gaps in the single-strand discontinuities are repaired by host cell enzymes (Engelman and Cherepanov, 2008; Engelman *et al*., 1991; Pauza, 1990).

Integration into chromosomal DNA is not the only destiny for viral DNA; as those viral DNAs that fail to integrate into the host genome can persist in the nucleus as episomal DNA, so called 'episomes'. Two types of episomes could persist as intermediate or by-product of viral integration: double or single LTR circles (see Figure 1.6). Double LTR circles are the result of either end to end joining, LTR to LTR, by non-homologous end joining (NHEJ), while single LTR circles the result of homologous recombination (HR) of one LTR to another creating an episome with only one LTR (Nakajima *et al*., 2001; Philpott and Thrasher, 2007). There is also the possibility of tandem joining of two episomes in auto-integration, which is non-functional (Farnet and Haseltine, 1991). The episomes are unable to replicate and they will be diluted with time in proliferating cells or degraded by cellular nucleases (Butler *et al*., 2002). Nevertheless, they are detected in the nucleus of transduced cells and are transcriptionally active (Banasik and McCray, 2009). Hence, if a therapeutic gene was present in the nucleus of a proliferating cell as an episome, it would not have long-term effect. Furthermore, the incorporation of a gene segment of a Scaffold/Matrix Attachment Region (S/MAR) from the human β-interferon gene, into vector transcription units to enhance transgene expression was achieved (Nielsen *et al*., 2009). However, this had only limited effects on the establishment of stable transgene in lentiviral vectors in contrast to, for instance SV40-

based vectors which has shown that controlled episomal replication could be stably maintained for over 100 generations (Piechaczek *et al*., 1999).

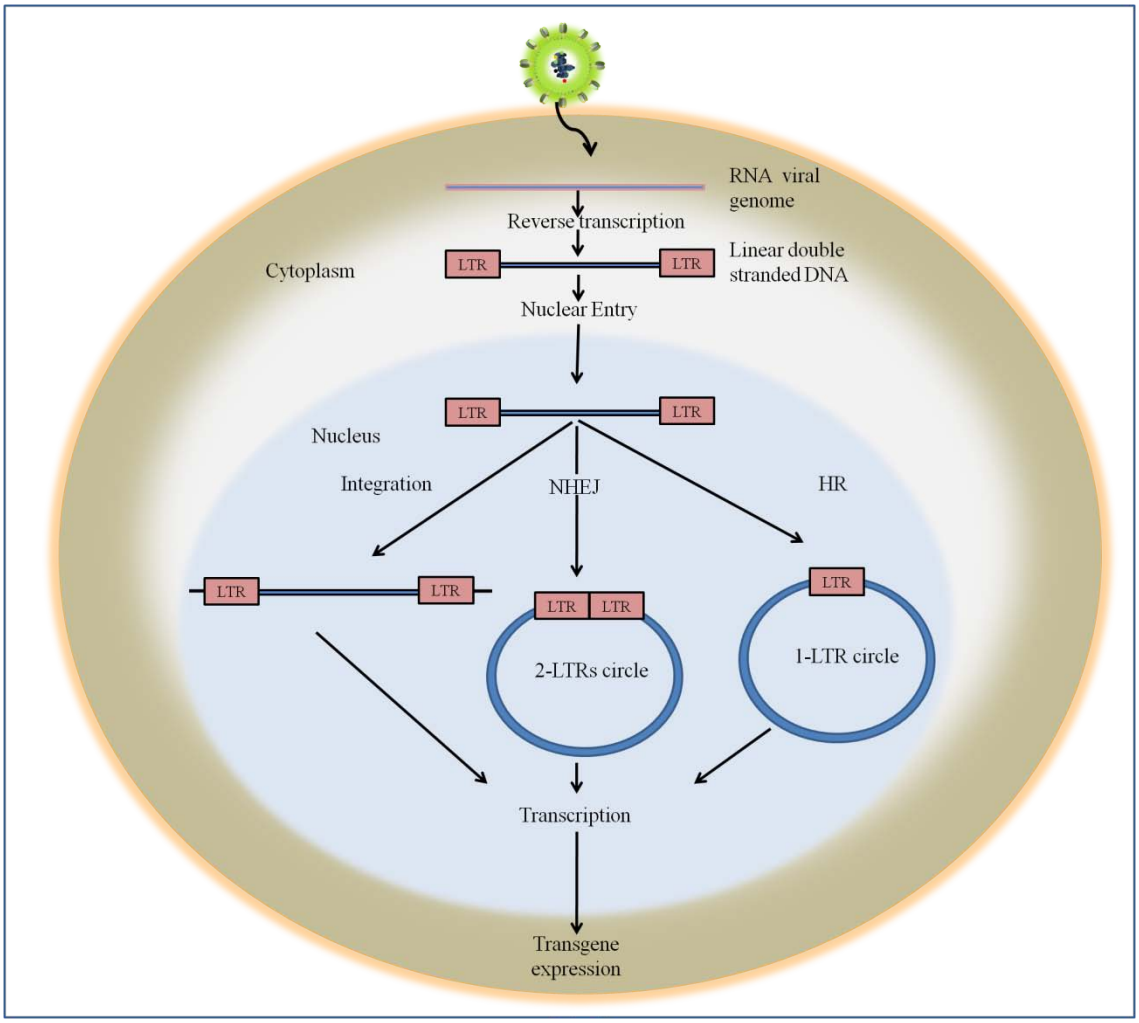



**Figure 1.6: HIV-1 based vector transcription model transcription.** Upon viral entry into the target cell, the viral RNA genome will be reverse transcribed in the cytoplasm into double stranded linear DNA. Subsequent nuclear entry will be facilitated by PIC leading to proviral DNA entry into the nucleus. Three possibilities could occur: integration of the proviral genome into chromosomal DNA, forming of 2 LTRs circle episomes via non-homologous end joining (NHEJ) or forming of 1 LTR circle episomes via homologous recombination (HR). Model adapted from Philpott and Thrasher 2007.

The natural tendency of LVs to form episome could be enhanced by impairing the virus's ability to integrate into chromosomal DNA avoiding all the related possibilities of mutagenesis and / or oncogenesis. This can be achieved by mutating the viral

integrase; accordingly there are two types of integrase mutations: class I and II mutations. Integrase class I mutations result from individual changes (point mutations) of specific amino acids leading to impaired viral integration and at the same time maintain the function of reverse transcription and nuclear import. On the contrary, class II mutations are usually occur in integrase active site leading to alteration of several functions of integrase which in turn will affect the integration step resulting in production of reduced amounts of viral DNA. Thus, this mutation related to complete integrase absence and faulty reverse transcription, nuclear localisation, nuclear entry and most importantly viral integration (Engelman, 1999; Lu *et al*., 2004).

The integrase active site encompasses three vital residues: aspartic acid at position 64 and 116 and a glutamic acid at position 152 (Leavitt *et al*., 1993). The so-called "D64V mutation" represent most common class I mutation. In this mutation the aspartic acid residue at position 64 is replaced by valine residue leading to generating integrase only faulty to catalyses viral integration while keeping its other functions (Engelman *et al*., 1995; Leavitt *et al*., 1996). It is noteworthy to mention that IDLVs possess, in general, lower expression levels than IPLVs in many cells types, while in other cells, like retinal pigment epithelium and neurons, there was no difference (Bayer *et al*., 2008; Yanez-Munoz *et al*., 2006). Because of their episomal nature, IDLVs will be diluted out in dividing cells as they are unable to replicate leading to limit their use into non-dividing cells (Wanisch and Yanez-Munoz, 2009).

## 1.3 Gene repair and DNA double-strand breaks

DNA repair mechanisms are essential for any living cell to maintain the structural integrity of the genome, which requires repair of any DNA damage with high fidelity. There are two types of DNA breaks: single-strand breaks (SSBs) and double-strand breaks (DSBs). The most hazardous DNA damages that cells could face are DNA DSBs, which can be caused by endogenous factors like collapsed replication forks or reactive oxygen species generated as a by-product of normal metabolic processes, and exogenous effects like exposure to ionizing radiation and certain chemicals (Mauro *et al*., 2012). Unrepaired DSBs may lead to very serious effects to the cells, like mutagenesis, tumorigenesis, genomic instability and cell death (Kirchgessner *et al*., 1995). In human cells, as an example, thousands of SSBs and DSBs events occur

naturally per day without necessarily leading to harmful consequences (Holmquist, 1998). Even so; a single DSB could potentially lead to cell death or oncogenic mutations in theory. Two main mechanisms are available for the cells to repair DSBs: homologous recombination (HR) which relying on recruiting an intact homologous segment of DNA, naturally from the sister-chromatid, as a template from which to copy the information across the DSB (Figure 1.7). An alternative pathway of DSB repair is non-homologous end joining (NHEJ), which non-specifically rejoins ends without homology and often results in short insertion or deletion (InDel) mutations at the site of break (McKinnon and Caldecott, 2007; San Filippo *et al*., 2008), or leading to chromosomal translocation (Brunet *et al*., 2009).

HR represents the major DSBs repair mechanism in bacteria and yeast but it is infrequent in the majority of higher eukaryotes (Johnson and Jasin, 2001). In contrast, it is believed that NHEJ mechanism play a dominant role in DSBs repair in higher animals (Iiizumi *et al*., 2008; Pruett-Miller *et al*., 2008). An emerging approach "gene repair" is introduced to gene therapy by harnessing the natural mechanism of HR to repair faulty genes at their endogenous loci maintaining genes integrity and regulation. Consequently, gene therapy strategies favor HR to modify the mutated gene rather than deliver a supplementary correct copy of the gene because HR mediates a genetic manipulation which ultimately restores a normal copy of damaged gene and guarantee permanent genetic alterations. Hence, an ideal correction would be achieved as the correct copy accurately replaces mutated sequences. Nevertheless, the low spontaneous frequency of HR in somatic mammalian cells has limited the therapeutic (and even experimental) gene modification strategies. In somatic mammalian cells (including human and murine cells) the expected spontaneous HR rate is (~ 1 in $10^6$ cells), while the rate of NHEJ is about (~ 1 in $10^3$ cells) (Gellhaus *et al*., 2010; Porteus and Baltimore, 2003; Pruett-Miller *et al*., 2008). For instance, in mouse embryonic stem cells the HR rate is about 1% of total random integration events (Osakabe *et al*., 2010). Thus, HR natural spontaneous frequency is too low to be used clinically for gene therapy studies; however this rate could be raised up to 20% when DSBs are introduced by nuclease-inducing DSBs like ZFNs in different cell types (Cornu *et al*., 2008).

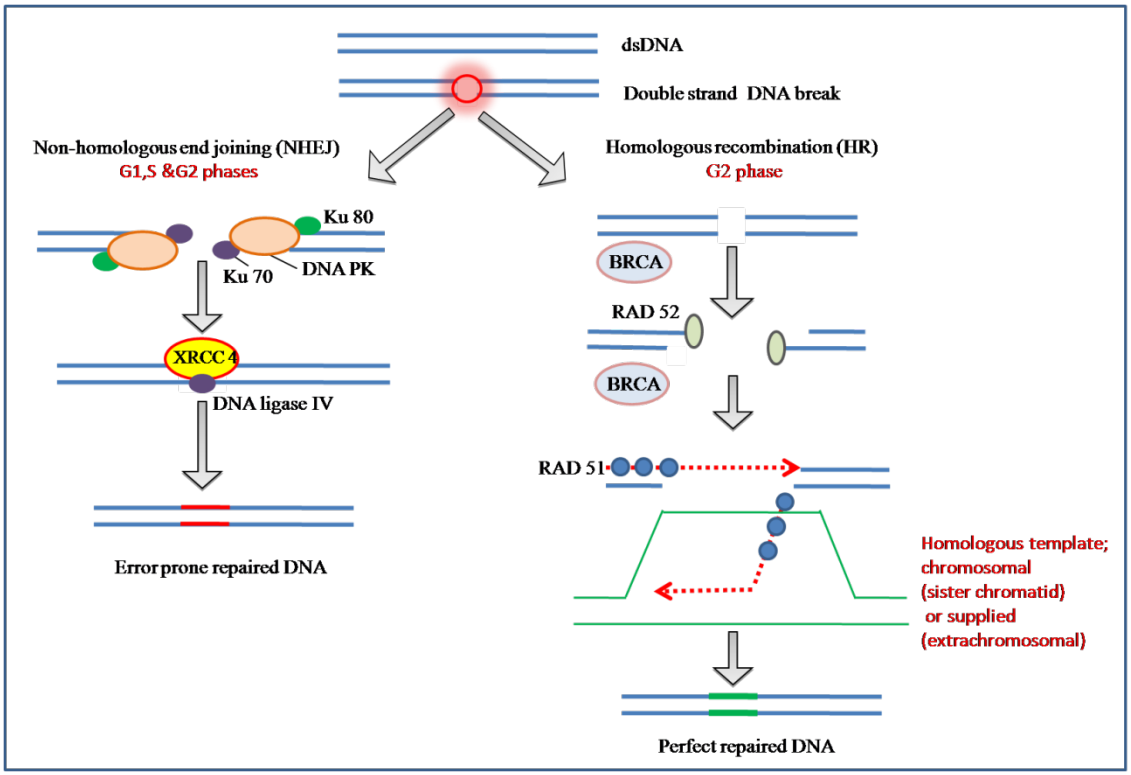



**Figure 1.7: Sschematic diagram of DSBs repair mechanisms.** Two main strategies are available in living cells to repair DSBs; the first is non homologous end-joining (NHEJ) which is considered as a dominant pathway in mammals and could be available during G1, S and G2 phases of cell cycle. During NHEJ, the broken strands of DNA are re-joined directly and non-specifically generating frequent mutations and eventually error prone repair. The DNA end-binding protein Ku 70 and Ku 80, recruits the DNA-dependent protein kinase (DNA-PK), XRCC 4, and DNA ligase IV for re-joining. While the second is homologous recombination (HR) mediated repair which is considered as a dominant pathway in lower organisms and could be available during G2 phase of cell cycle. BRCA co-localizes with RAD 51, which is implicated in DSB repair through its ability to bind single-stranded DNA and facilitate DNA strand exchange. Then, Rad 52 binds to the resected ends, and Rad 51 forms a filament that facilitates DNA strand invasion. The resected 3'end invades a homologous DNA duplex and is extended by DNA polymerase. HR naturally uses a sister chromatid homologous template or any other supplied extrachromosomal homology generating a perfect repair. The diagram was amalgamated from San Filippo *et al.* 2008, and Weberpals *et al*. 2008.

#### 1.3.1.1 The role of HR in gene repair

In 1978, the concept of homologous gene targeting was first described in yeast by Hinnen *et al*. (Hinnen *et al*., 1978), then it is successfully transferred to mammalian cells by Mario Capecchi, Martin Evans and Oliver Smithies. For that, Capecchi, Evans and Smithies have jointly awarded the Nobel Prize in physiology or medicine for 2007 for their discoveries of "principles for introducing specific gene modifications in mice by the use of embryonic stem cells" (http://www.nobelprize.org/nobel_prizes/medicine/laureates/2007/press.html, accessed on 4/7/2012). They opened the door for gene repair approach by their recognisable achievement of demonstration of that mammal genes could be targeted by HR in cultured cells and changes transferred into the mouse germ line using embryonic stem cells. Their pioneering research has highlighted the importance of therapeutic gene repair because it has made the cell genetically wild type, restored the function of faulty gene permanently and avoided the insertional mutagenesis.

During the process of DSBs repair by HR, a second homologous copy to the mutated sequence must be available to act as a template. The cell usually prefers to use the sister chromatid as a donor-template for DNA polymerase to extend and repair the DSBs. Thus, cell cycle phase could be the major determinant for the genomic DSB to be repaired by NHEJ or HR (Olsen *et al*., 2010). As a result HR is most active at late S and G2 phases of the cell cycle, while NHEJ can be active at all the phases (Kotnis *et al*., 2009), but it is most essential in G1 and G0 phases (Wiegant *et al*., 2010). As an example for this preference *Saccharomyces cerevisiae* prefers to use the sister chromatid by 2-3 orders of magnitude over a homologous or heterologous chromosome (Johnson and Jasin, 2001).

Interestingly, the cellular machinery controlling HR could be exploited for gene repair by providing an extrachromosomal homologous sequence to act as a template to replace precisely the mutant one but this exogenous DNA contains mismatches or intervening sequences. Of note, HR machinery can be stimulated by 2 to 3 orders of magnitude when homology is provided either from extrachromosomal DNA in gene repair experiments or from a repeated chromosomal sequence (Elliott *et al*., 1998). The two DNA ends resulted from DSB (the broken ends) will be under the influence of exonuclease action, generating single stranded 3’ ends in both sequences. These single

stranded 3’ ends invades the donor DNA template (whether a sister chromatid or extrachromosomal DNA) at the homologous region, producing DNA duplexes with a donor template. Later on the DNA polymerase extends the 3’ ends using the donor DNA as a template (Wong and Chiu, 2011). At the end, HR will mediate integration of the new sequence into a small fraction of the targeted cell genome; and if selector genes are included in the donor sequence (to be removed eventually), the genetically modified cells can be isolated (Capecchi, 2001). Furthermore, it has been confirmed recently that arresting the targeted cells at the late S and G2 phases of the cell cycle would increase the HR rate by up to 5 fold (Connelly *et al*., 2010; Olsen *et al*., 2010), simultaneously that will not have any influence on the activity of NHEJ repair pathway (Guirouilh-Barbat *et al*., 2008).

There are two conserved class of enzymes, so called recombinases, control the HR function (San Filippo *et al*., 2008). Those recombinases include the RAD51 group (containing RAD50, RAD51, RAD52, and RAD54) and Dmc1 group (containing RPA, XRCC2, XRCC3, and BRCA) (Johnson and Jasin, 2001). Both recombinases mediate the formation of DNA joints that link homologous DNA molecules, but they are active at different conditions. The Rad51 enzyme is active in somatic and meiotic cells, while the second recombinase, Dmc1, is active only in meiosis and acts in concert with Rad51(San Filippo *et al*., 2008). The proteins controlling the NHEJ activity will be discussed below (section 1.8.2). It has been suggested that there is competition between NHEJ and HR mechanisms upon the DSBs repair (Iiizumi *et al*., 2008; Ramirez *et al*., 2012). However; the selection is more probably to be decided by the functional proteins that initially recognize and process the DNA ends at the DSBs locus (Yun and Hiom, 2009).

The efficiency of HR mediated gene repair depends on different factors:

- delivering of intact homologous template into the nucleus of targeted cells.
- the efficiency of transfection or transduction efficiency of targeted cells.
- the nature and type of cell ; it is essential to understand and modify the balance of HR and NHEJ mechanisms.
- the type of genetic mutation to be repaired; it would be ideal therapy if the repaired cells possess a selective advantage as for instance in some SCID disorders.

- the induction of genome site specific DSBs to stimulate HR (Ellis *et al*., 2012; Urnov *et al*., 2010).

It has been reported that HR-mediated manipulation could be possible at least in hematopoietic stem cell (HSC) progenitors (Hatada *et al*., 2000) because of many reasons: the repaired cells would possess a selective repopulation advantage, the suitability of HSCs to many viral vectors including (lentivirus based vector) and the ease of *ex vivo* gene therapy (Cartier *et al*., 2012; Langford-Smith *et al*., 2012). On the other hand, this approach is hindered by two main problems: the low abundance of these cells and the low achieved HR frequency. However; several procedures can be adapted to improve the HR frequency such as:

- increasing the length of the homologous template, for example it has been found that doubling the length will lead to a 10 fold increase in the gene repair frequency (Deng and Capecchi, 1992),

- pretreatment of the HSCs with appropriate cytokines will lead to 10 times increase in their repopulating abilities (Glimm *et al*., 2000),and

- it has been recently shown that introduction of a DSB at a specific genomic target locus can increase targeting frequencies (Carroll, 2008).

## 1.4 Induction of DSBs

As mentioned above, HR is inefficient in mammalian cell gene repair therapies because of: the low spontaneous frequency and the vectors integrate randomly into the host genome sites via NHEJ at least 3000 times more (Iiizumi *et al*., 2008). However; there are several ways to improve the efficiency of gene repair including: the use of recombinant AAVs (Pruett-Miller *et al*., 2008), and gene repair stimulation by inducing of DSBs at the target sites. It has long been confirmed that the induction of site-specific DSBs in the target sequence raises the HR rates by up to 3000 fold (Olsen *et al*., 2010). Additionally, the presence of genomic DSB inducers not only enhance the DNA damage response signalling pathways but also changes the activity of many central cellular proteins, affect the cell cycle and impinge on cell death (Harper and Elledge, 2007). Furthermore, it has been reported that induction of DSBs will not only enhance the repair efficiency but also the range of possible types of corrections resulting for

instance to large insertions of 8 kb (Moehle *et al*., 2007). Thus, the new emerging approach of "genome surgery" has revolutionized gene repair research using custom-designed enzymes directed to introduce genome site-specific DSBs. These designer nucleases include: ZFNs, meganucleases, and transcription activator-like effector nucleases (TALENs).

### 1.4.1 ZFNs

#### 1.4.1.1 ZFN structure and function

ZFNs are engineered chimeric proteins that link wild-type *Fok* I endonuclease domains to zinc finger DNA binding domains (Kandavelou *et al*., 2005; Porteus and Carroll, 2005). These proteins are custom designed to induce DSBs at very specific DNA sequences (Urnov *et al*., 2005). Each ZFN protein consists of two monomers and typically every monomer incorporates a non-specific cleavage domain of the *Fok* I type II restriction enzyme (a restriction endonuclease from *Flavobacterium okeanokoites*) (Liu *et al*., 2010) fused to about 3-4 zinc finger domains (Durai *et al*., 2005; Liu *et al*., 2010). Each zinc finger domain is particularly designed to recognize 3-4 base pairs (bp) of double stranded DNA. Normally, three (Kandavelou *et al*., 2005) or four (Urnov *et al*., 2005) zinc domains are linked together in tandem to generate a zinc finger protein (ZFP) that recognizes and binds to a 9 bp (Kandavelou *et al*., 2005) or 12 bp (Urnov *et al*., 2005) respectively, depending on the basic design for the ZFN. This 9 or 12 bp of the DNA sequence which can be recognized by one ZFP is known as the ZFN (or monomer) binding site, with one binding site for each monomer (see Figure 1.8).

Historically, the zinc finger domains were first discovered in *Xenopus laevis* (Durai *et al*., 2005), later it became known that they are naturally present in the genomes of different organisms including human (Klug, 2005). There is a high degree of conserved residues within the fingers; however there are some differences in few residues allowing them to bind specifically to distinct DNA sequences. Multiple fingers are assembled in a modular tandem fashion to recognize a specific DNA sequence (Cathomen and Weitzman, 2005). The three dimensional structure of the zinc finger has been shown to be composed of two β-sheets in an antiparallel configuration, which forms a loop produced by the two cysteines and an α-helix including two histidines loop (see Figure

1.9 a). The two structural entities are stabilized together by the zinc ion and for this reason this structure is called as ββα motif (Klug, 2005).

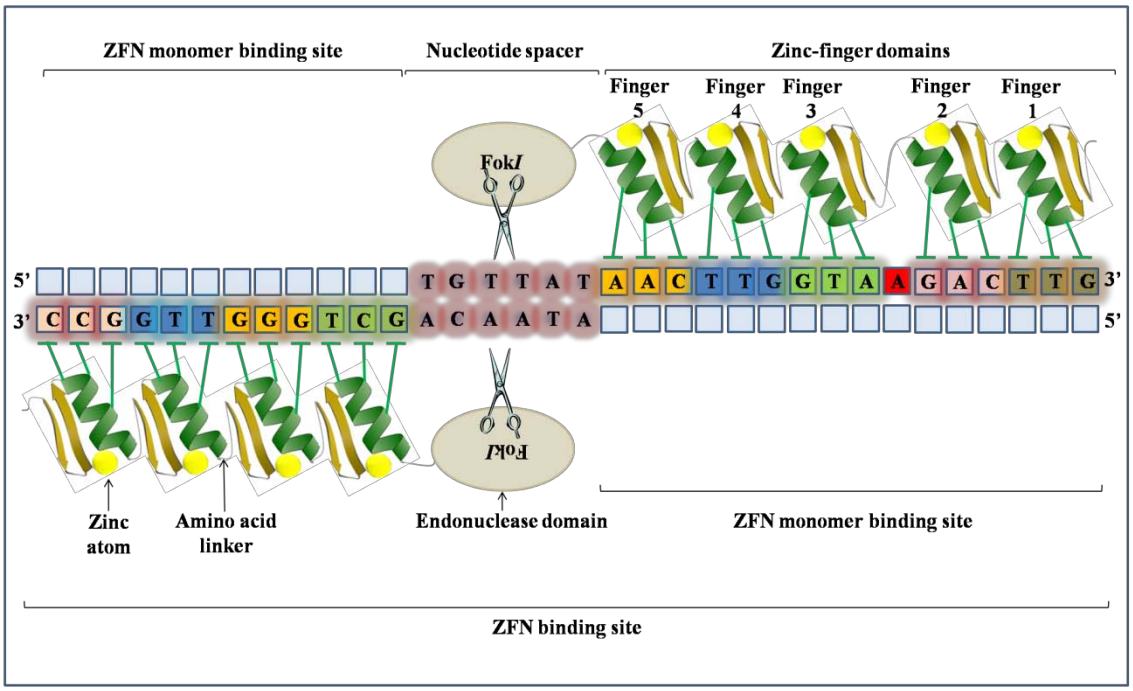



**Figure 1.8**: **Schematic representation of ZFN structure targeted against mouse *Prkdc***. ZFNs are dimeric proteins with 5 zinc-finger domains forming the right ZFN monomer binding site or 4 zinc-finger domains forming the left monomer. Zinc-finger domains are fused to the endonuclease domain of *Fok* I; each finger recognizes 3 base pairs of double stranded DNA via a single α -helix. The binding sites for the monomers are separated by a nucleotide spacer where DSB takes place.

The two ZFN binding sites are separated by approximately a 5-7 bp spacer, where the DSB occurs (Townsend *et al*., 2009). Altogether, the two ZFN binding sites (18 or 24 bp, in addition to the nucleotide spacer in between) represent the full ZFN recognition site (Dolan, 2006). So, the ZFN full site is long enough to be a unique locus within mammalian genomes (Kandavelou *et al*., 2009). The *Fok* I domain is able to dimerize, inducing a sequence specific DSBs whenever the two ZFPs bind their equivalent DNA target sequence in the opposite orientation (Zou *et al*., 2009). Hence in order for the ZFNs to create the cutting at the desired sequence, two *Fok* I must be dimerized. In fact, the actual DSBs of the targeted DNA sequence is initiated by the *Fok* I domain, a fact that confirmed by Pruett-Miller *et al* when HEK 293 cells were transfected with

designed ZFNs lacking the *Fok* I domain; there were on detected cytotoxicity (Pruett-Miller *et al*., 2009). It is not essential for the two monomer recognition sites to possess the same length to provoke specific targeted DSBs. By modification of the zinc domains, ZFNs can be precisely directed to induce DSBs at any region within the genome using the zinc finger targeter (ZiFiT) software that predicts possible genomic ZFN binding sites for engineered ZFPs (Osakabe *et al*., 2010; Sander *et al*., 2010).

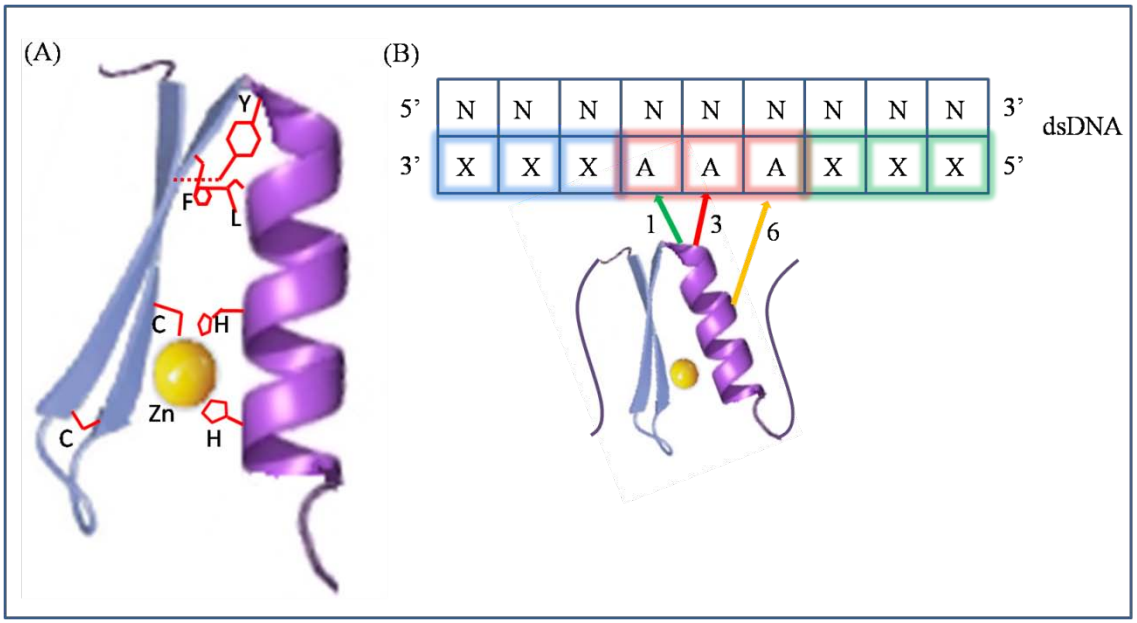



**Figure 1.9: The basic structure of zinc finger domain.** Panel (A) antiparallel β-sheets of two cysteines and an α-helix of two histidines forming the ββα structure, which is stabilized by zinc atom. Panel (B) is showing the primary attachments between α-helical amino acids at the positions 1, 3 & 6 (green, red & yellow arrows respectively) and the cognate DNA triplet (letters A). C: cysteine, H: histidine, F: phenylalanine, Y: tyrosine, L: Lucien, Zn: zinc atom, triple A: recognized DNA sequence by single zinc finger domain. Amalgamated from Dolan 2006, and Klug 2005.

The basic structure of a zinc finger consists of 30 amino acids within the ββα structure (Dolan, 2006). Predominantly, there are two cysteine and two histidine residues in every single zinc finger, in addition to other three dominant amino acids, tyrosine, leucine and phenylalanine; therefore this domain is also known as the C2H2 domain. It has been suggested that these seven amino acids probably form a hydrophobic structural centre in the zinc fingers; as well as providing a vital role in determining the tertiary folding configuration of the zinc fingers; see Figure 1.9. The primary interactions

between the zinc fingers and the DNA bases are produced by the α-helix. The α-helical amino acids at positions 1, 3 and 6 contact the three (triplet) consecutive DNA sequence bases by binding in DNA major groove via particular hydrogen-bond interactions (Dolan, 2006). Apparently, the mechanism of zinc finger recognition to the DNA target sequence is a one-to-one contact between particular α-helical amino acid to particular DNA nucleotide base (Figure 1.9 b). This mode of interaction differs from other DNA-binding proteins, where one amino acid contacts two nucleotides or vice versa (Klug, 2005).

Klug has reviewed a basic technique to produce target site-specific ZFN domains. The ZFN protein domains are cloned as fusions to the minor coat protein (pIII) of bacteriophage fd, which in turn, initiates their expression on the tip of the viral capsid. Phage showing the peptides of interest can be affinity purified by binding to the target and then amplified in bacteria for use in further rounds of selection and for DNA sequencing of the cloned genes (Klug, 2005).

ZFN's mode of action is adapting the recombinogenic repair mechanism of DSBs in the target DNA sequence of the living cells (Shukla *et al*., 2009). When site specific DSBs are induced by the use of ZFNs, the cellular DNA damage response will be stimulated leading to repair these DSBs by either NHEJ or HR. Importantly, it has been suggested that ZFN cutting could take place mainly during S phase of the cell cycle, when all genomic DNA initialise for replication (Carroll, 2011). Consequently, by harnessing the cell's individual DSBs repair mechanism, specific gene alterations can be generated. As a result, ZFNs have been developed to induce precise genetic modifications in many different types of living cells. This implies that, hypothetically at least, it is possible to target any gene in the genome, through the manipulation and recombination of ZFPs with varying specificities (Porteus, 2009). Recently, the ZFNs technique has been extensively utilized for genome modifications at endogenous loci in eukaryotic systems; since the zinc-finger domain can be designed to recognize unique DNA sequences (Shukla *et al*., 2009). It has became clear that using ZFNs will induce site specific DSBs enhancing gene targeting at precise loci in human cells (Hockemeyer *et al*., 2009; Zou *et al*., 2009), plant models (Shukla *et al*., 2009; Townsend *et al*., 2009) and animal model (Li *et al*., 2011a).

#### 1.4.1.2 ZFN applications

ZFN technology has been successfully used; so far, for:

- repression of gene expression; for example in a mouse cell line by selective targeting for genetic alterations in pathogens or transformed cells and in HIV replication (Reynolds *et al*., 2003),
- regulation of promoter specific activation by switching on/ off gene expression; for example in herpes simplex virus (Papworth *et al*., 2003),
- gene knocking out by sequence-specific ZFN mediated NHEJ (Geurts *et al*., 2009a; Malphettes *et al*., 2010),and
- gene repair by sequence-specific ZFN mediated HR (Olsen *et al*., 2009).

From practical point of view, it has been shown that site specific ZFNs can significantly increase the gene targeting frequency several orders of magnitude. For example, Kandavelou *et al* has pointed out that gene correction using ZFN was about 200 fold more efficient in albino mouse melanocytes (Kandavelou *et al*., 2009). Moreover, Zou *et al* (Zou *et al*., 2009), has demonstrated that the use of site specific ZFN improved the HR frequency 2400 fold in human embryonic stem cells. Impressively, Townsend *et al* confirmed that gene targeting frequency exceeded 2% in plant cells (Townsend *et al*., 2009). While using ZFNs in HEK 293T or human chronic myelogenous leukemia K-562 cells frequencies ranged between 2 and 20% (highest in cells temporarily arrested at G2/M boundary and targeting the endogenous *IL2RG* gene rather than a chimeric eGFP target locus); importantly, biallelic correction was detected in up to 6.6% of cells. Gene targeting of *IL2RG* in primary CD4+ T cells occurred in 5.3% of cells (Urnov *et al*., 2005). Alternatively, gene disruption can also be obtained at high frequency (13%) in human stem cell lines by NHEJ repair of a ZFN-induced DSB (Lombardo *et al*., 2007). This *CCR5* knockout strategy is paradigmatic for the generation of CD4+ T cells resistant to HIV-1 infection (Perez *et al*., 2008). These impressive results and other (de Pater et al., 2009; Perez et al., 2008) indicate that the specificity of the designer ZFNs majorly improves the efficacy and the efficiency of the gene targeting rates.

Increasingly, ZFN technology has been exploited in gene targeting to provide DSBs at specific DNA sequences, a process that many researchers (Carroll, 2008; Kandavelou *et al*., 2005; Lombardo *et al*., 2007; Porteus and Carroll, 2005) have adapted to enhance

the HR mechanism in gene repair. The impressive results, higher modification frequencies and precision of ZFN-driven HR could revolutionise the field of gene therapy for mammalian somatic cell genetics and the study of gene function. This study will adapt the same strategy to enhance gene repair by HR using especially designed ZFNs. This new designed ZFN has been recently engineered, particularly for repairing the mutant *Prkdc* gene, by Sangamo BioScience Inc., Dr Yáñez's collaborators in somatic gene therapy.

#### 1.4.1.3 ZFN-related toxicity

In spite of successful ZFN applications; ZFNs exhibit no perfect action. A major limitation affecting the application of ZFN technology is the related cytotoxicity; which has been reported in human cells, zebrafish and *Drosophila* (Zhang *et al*., 2010b). Related toxicity can be attributed to either the high expression of the ZFNs leading to induce more DSBs (de Pater *et al*., 2009), or the off-target DNA induced DSBs at related sequences somewhere else in the genome because of the non-specific binding. These unwanted DSBs may possibly result in non-anticipated genotoxic effects. Some of these genotoxic problems could be seriously adverse to the cell, generating chromosomal abnormalities as the unpredictable DSBs could result in chromosomal translocations or NHEJ (Dolan, 2006); and eventually to cell death and elimination of targeted cells. Other genotoxic effects may cause minor alterations and induce less harmful changes to the cells such as small deletion and addition mutations that could be continued in the existing population (Pruett-Miller *et al*., 2009; Zou *et al*., 2009). Additionally, NHEJ mediated repair could lead to introduce undesirable alterations at other off-target genomic sites; it has been reported recently that ZFNs introduce a wider range of off-target DSBs (and as a result NHEJ-mediated mutations) than previously expected (Gabriel *et al*., 2011; Pattanayak *et al*., 2011).

One possibility to minimize the genotoxic effects is to improve the specificity of the recognition of the designer ZFNs for their target locus; this can be achieved by engineering more than 3 zinc finger domains per monomer. Accordingly, the length of the targeted DNA will be increased, thus exponentially maximising the level of uniqueness. A good example for this successful strategy was reported by Hockemeyer *et al*, when ZFNs were designed to contain 4-6 zinc finger domains which collectively bind to DNA stretch of 24-36 bp to guarantee a high degree of rarity within the targeted

sequence (Hockemeyer *et al*., 2009; Porteus, 2009). In another study, ZFNs with 5-6 zinc fingers recognizing 30-36 bp of the DNA sequence have been engineered to maximize the specificity and elongate ZFN recognition site (Geurts *et al*., 2009a). More recently, ZFNs with 5 zinc domains in one monomer and 6 domains in the other, have been designed to recognise 38 bp of the targeted locus (Malphettes *et al*., 2010).

The second promising strategy to reduce the ZFN off-target cleavage is to re-engineer the *Fok* I domains in a way that preventing, or at least minimising, the homodimerization of ZFNs. As mentioned above, a pair of *Fok* I domains must dimerize to activate the ZFN mechanism, which eventually will induce site specific DSBs. It seems that *Fok* I domains are naturally able, even if they are not bound to any DNA sequence, to create homodimers although feebly (Mani *et al*., 2005). In addition to that, normal dimerization of *Fok* I domains is quite weak (Miller *et al*., 2007). It has been shown that the *Fok* I domains could be altered such that only heterodimeric ZFNs could form upon binding to DNA leading to significantly reduced ZFN-associated cytotoxicity (Sollu *et al*., 2010). Therefore, the engineering of obligatory heterodimers of the *Fok* I domains is vital to avoid or at least to significantly decrease the levels of genotoxicity in human cells (Kandavelou *et al*., 2009). As a result, this will ensure that both left and right monomers are situated at their proper target recognition sites on the DNA sequence.

It has been reported that longer term of ZFN expression could afford extra time for off-target DNA cleavages leading to further alterations and cytotoxic effects (Remy *et al*., 2009). Consequently, an alternative strategy to reduce the cytotoxicity of the ZFNs has been applied by limiting ZFNs expression. There are several approaches have been reported here:

- attenuation of ZFNs toxicity through regulation of protein levels, which in turn creates ZFNs with shortened half-lives. This is achieved by engineering small molecules which destabilise ZFNs leading to regulate protein levels, this approach blends both high rates of ZFN-mediated gene targeting while reducing ZFN toxicity (Pruett-Miller *et al*., 2009).
- direct injection of ZFNs as RNA molecules into the targeted cell (Mashimo *et al*., 2010). This may be additionally favourable because of: (i) apparently, RNA has shortened half-lives; therefore the ZFNs will be present for only limited duration

(Beumer *et al*., 2008; Zhang *et al*., 2010b), (ii) remarkably, there will be an acceleration of the expression rate, and (iii) reduced possibility of ZFN insertional mutagenesis, as RNA cannot integrate into genomic DNA.

- transient DNA delivery (Zhang *et al*., 2010b).
- inducible expression systems (Zhang *et al*., 2010b).

Another possible strategy to improve ZFN efficiency and reduce off-target cuttings has been tested recently by using mild hypothermia (cold shock by incubation of the ZFN transfected cells at $30^0$C) leading to several fold increase in ZFN activity. ZFN-hypothermia has generated a proficient and constant genome editing in different cells which can be attributed to the extensive reduction in mRNA and protein turnover, initiation of molecular chaperones, or viability enhancement and delayed apoptosis (Doyon *et al*., 2010).

For current study, two important improvements in ZFN technology will be adapted to minimize possible genotoxicity. The first will be the use of reengineered ZFN *Fok* I domains to avoid homodimerization by using of obligatory heterodimers *Fok I* domains. This step has been done by our collaborator; Sangamo BioScience Inc (California, USA). The second improvement will be via controlling ZFN delivery system. The ZFNs genes will be delivered by incorporation them into IPLVs and IDLVs. For clinical application it is hoped that IDLVs could be efficient to deliver ZFNs at levels enough to induce DSBs, as the use of IDLVs will decrease the amount of vector integration considerably. Despite the fact that the use of IDLVs will not eliminate such integration events absolutely; this approach will rely on their short term expression to avoid permanent integration because of their episomal nature. Hence, this approach will not allow additional time for ZFNs to create unpredictable off-target cleavages.

### 1.4.2 Meganucleases: the homing endonucleases

Another technology to introduce site specific DSBs and maybe exploited in HR leading to gene repair, is the use of meganucleases. Meganucleases are a varied group of highly specific endonucleases, which recognize a large stretch (12-45 bp) of DNA sequence (Epinat *et al*., 2003; Moure *et al*., 2002). These nuclease could be found inserted in different genomes in eukaryotes, prokaryotes and viruses (Moure *et al*., 2002). In nature, meganucleases are embodied by “homing endonucleases”; a large family of

DNA endonucleases. Meganucleases are generally encoded within introns or inteins (segments of proteins that are able remove themselves and rejoin the remaining portions) although some exceptions are also present (Grabher and Wittbrodt, 2007). They called “homing” because they mediate the gene mobility event of their hosting genes (Moure *et al*., 2003), resulting in DSBs at the targeted sequences which do not include these mobile elements (Moure *et al*., 2002). Site-specific DSBs induced by meganucleases in living mammalian cells increased gene repair rates by a number of orders of magnitude (Epinat *et al*., 2003).

Homing endonucleases are divided into five families based on sequence and structure motifs: LAGLIDADG, GIY-YIG, HNH, His-Cys box and PD-(D/E)XK (Orlowski *et al*., 2007; Zhao *et al*., 2007). The largest one is LAGLIDADG family and hold either one or two sets of the LAGLIDADG sequence motif. Endonucleases with one motif act as homodimers like I-*Ceu*I or I-*Cre*I; while those enzymes with two LAGLIDADG motifs act as monomers like I-*Sce*I, PI-*Sce*I and I-*Dmo*I (Duan *et al*., 1997; Heath *et al*., 1997). In all members of LAGLIDADG family, the LAGLIDADG motif generates two full α-helices where a 2-fold symmetry axis separates two monomers or apparent domains. On both sides of the LAGLIDADG α-helices, there are a four-stranded β-sheet offering a DNA recognition part which mediates the binding between endonucleases and a half sequence of the recognition site of the target DNA locus, see Figure 1.10 (Epinat *et al*., 2003).

Usually, meganucleases are acting as movable DNA elements and they are able to bind to DNA sequences which match to intron-free or intein-free genes, where DSBs will be introduced. Those DSBs will be repaired ultimately by cellular repair machinery leading to insertion of the intron or intein where the DSBs induced (Epinat *et al*., 2003). Therefore, organisms encoding homing endonucleases are not affected by their own endonucleases as they located on either intron or intein which splice them from the host mRNA or protein, respectively (Moure *et al*., 2002).

Homing endonucleases are similar to restriction enzymes in their ability to induce DNA DSBs precisely in the presence of divalent cations; however there are several differences between them: (i) the recognition site of the restriction enzymes is shorter (3-8 bp) than homing endonucleases one; (ii) the restriction enzymes are extremely sensitive to any single nucleotide alteration in their recognition site, whereas homing

endonucleases are more rather tolerant of single nucleotide alteration; (iii) homing endonucleases mostly fall within one of four families, while restriction enzymes do not subdivided into identifiable families; and (iv) they differ in their genomic location (Belfort and Roberts, 1997).

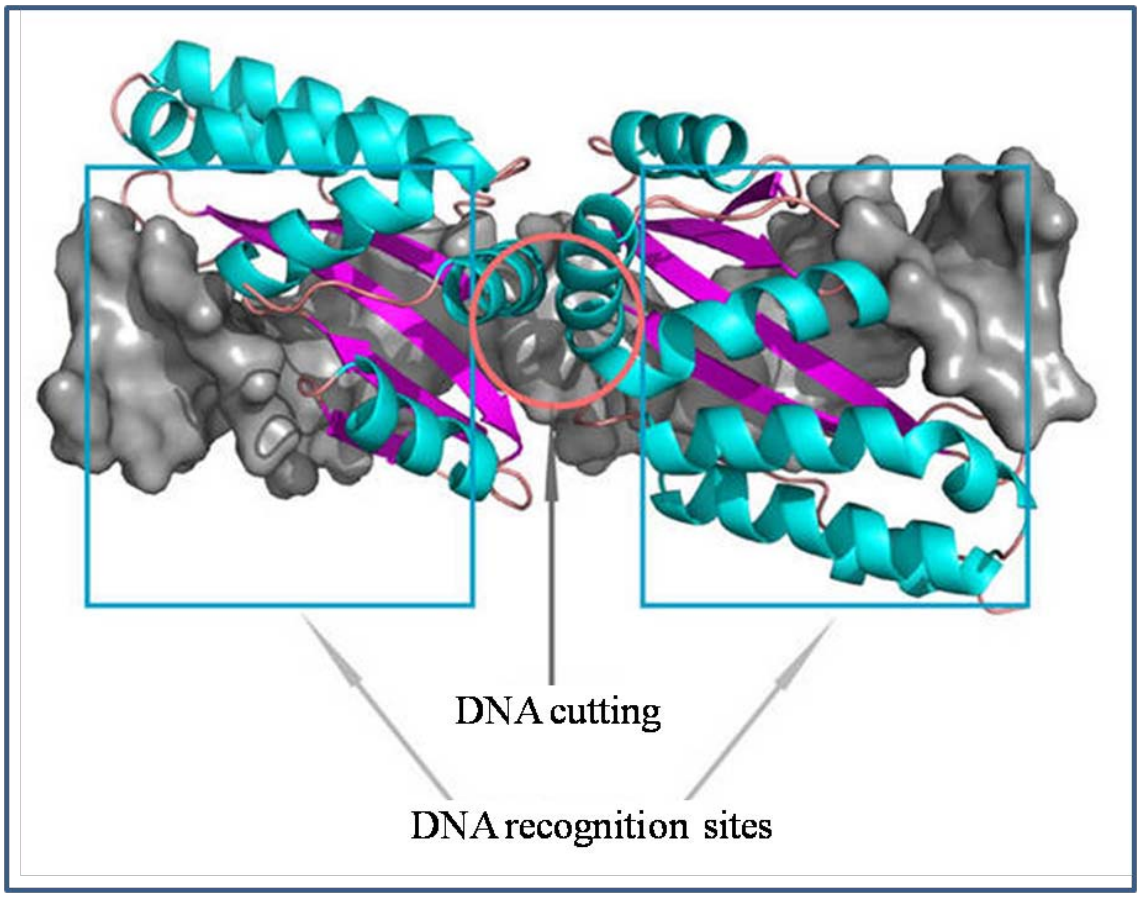



**Figure 1.10: Schematic representation of monomeric structure of homing endonuclease.** In monomeric enzyme, each monomer has two copies of LAGLIDADG motif forming four α-helices in total (blue) and four β-sheets (purple). The induced DSBs will happen at the centre of DNA recognition site. Reprinted from http://www.cellectis.com/genome-engineering/meganucleases/engineered-meganuclease/meganuclease-technologies/. Accessed on 20/06/2012.

Homing endonucleases bind their DNA targets within the recognition site via many direct phosphate and base-specific contacts, which are equally distributed between the two domains, in addition to hydrophilic interactions which represent a good indicator of the high specificity of these enzymes (Moure *et al*., 2003). However, they might

introduce random DSBs in their target DNA due to off-target cutting (Moure *et al*., 2002). As a result, the specificity of DSBs at defined sites expanded the use of meganucleases in gene repair strategies in different fungal, plant and mammalian applications (Belfort and Roberts, 1997). The use of majority of natural meganucleases in gene repair studies relies upon the presence of enzymes whose recognition sites are previously encoded in the target sequence. However, this occurrence is very rare since the usual collection of homing endonucleases, at present, is restricted to about 300 enzymes, the majority of them still uncharacterized or theoretical. Currently, meganucleases have become more attractive because of continued engineering efforts that generated custom designed meganucleases targeting several genomic sequences. These modified proteins are based, for example, on subunit inter-changeability by fusing both similar and different α / β domains, and some mutations in specific residues (Silva *et al*., 2011).

I-*Sce*I, from *Saccharomyces cerevisiae*, was the first discovered homing endonucleases (Moure *et al*., 2003). This enzyme belongs to the LAGLIDADG family and it is an intron-encoded endonucleases. The I-*Sce*I endonucleases acts as a monomeric form of 235 amino acids, including two conserved motifs (the two LAGLIDADG motifs) and it is catalysed by either magnesium or manganese (Grabher and Wittbrodt, 2007). DSBs result from I-*Sce*I action will have a 4 bp overhang presenting a 3'-hydroxyl terminus (Moure *et al*., 2003). The recognition site of the I-*Sce*I endonuclease is 18 bp long (TAGGGATAACAGGGTAAT), which is very rare in genomic DNA because of its theoretically low frequency ,only one time in $7x10^{10}$ bp of random sequence (Choulika et al., 1995a; Thermes et al., 2002), or once per 20 human genomes (http://www.cellectis.com/genome-engineering/meganucleases/engineered-meganucleases/meganuclease-technologies/, accessed on 20/06/2012). Additionally, pre-engineered I-*Sce*I sites can tolerate quite a few nucleotide alterations, making this endonuclease the enzyme of interest to induce site specific DSBs at selected genes in yeast, plants, insects, fish, and mammalian applications (Moure *et al*., 2003). Finally, vertebrate genomes examined so far are free from I-*Sce*I sites, eliminating the risk for genome fractionation by the action of meganucleas (Grabher and Wittbrodt, 2007).

It is worth mentioning that designing of meganuclease novel DNA binding without disrupting their catalytic activity it is confirmed to be quite challenging as their DNA

recognition and cleavage domains are not structurally independent as they are for ZFNs and TALENs (Ramirez *et al*., 2012). Two research groups had reported that co-transfecting a plasmid expressing I-*Sce*I together with the targeting construct in cell lines led to up to 4-log increases in gene repair frequencies (Choulika et al., 1995b, a; Smih et al., 1995). Additionally, I-*Sce*I vector dose-dependent and cell type-dependent targeting frequencies were noticed, being about 5-fold higher in U-2 OS (human osteosarcoma) and HEK 293T cells than in HT1080 (fibrosarcoma) cells (Cornu and Cathomen, 2007). Moreover, high-efficiency gene conversion has been demonstrated with chimeric meganucleases at chromosomal target loci. Frequencies approached 0.2% with a designer meganuclease in Chinese hamster ovary (CHO) cells (Arnould *et al*., 2007). Also, it has been reported that engineered meganucleases have successfully cleaved two different human genes, XPC and RAG1 (Arnould *et al*., 2007; Smith *et al*., 2006).

### 1.4.3 TALENs

Genome surgery can also be achieved by harnessing the transcription activator-like effector (TALEs) nucleases to introduce site specific DSBs. The TALE modules a family of proteins used in the infection process by plant pathogenic bacteria *Xanthomonas oryzae*, (Boch, 2011; Sugio *et al*., 2007). TALE nucleases (TALENs) are DNA recognition repeats of native or synthesized transcription activator-like fused to DNA cleavage domains of *Fok* I. These customized proteins are synthesized using a modular assembly method (Li *et al*., 2012a). Their DNA binding domain are composed of so- called TALE repeats, which they are ~34-35 multiple amino acids arranging in tandem. The basic structure for the TALEN is showed in Figure 1.11.

The structural features and domains of each TALE are composed of: an N-terminal translocation region, a central repeat region which recognises target DNA sequence, a C-terminal region which comprises NLS and an acidic activation domain. These nuclease are differ generally in the number and order of their TALE repeats in the DNA recognition domain (Mussolino *et al*., 2011). The TALE repeats sequence is almost identical apart from two highly variable amino acids in positions 12 and 13 of each repeat comprising a repeat variable di-residue (RVD). The RVD of each domain represents the base of recognition site for each unit to recognize one nucleotide in the

target DNA sequence, thus each individual domain can recognise one DNA base pair in the TALE recognition sequence (Boch *et al*., 2009; Hockemeyer *et al*., 2011; Moscou and Bogdanove, 2009). TALENs are recently used to induce site-specific genomic modifications in plant cells (Mahfouz *et al*., 2011), animal (Sander *et al*., 2011), and even in human pluripotent cells (Hockemeyer *et al*., 2011).

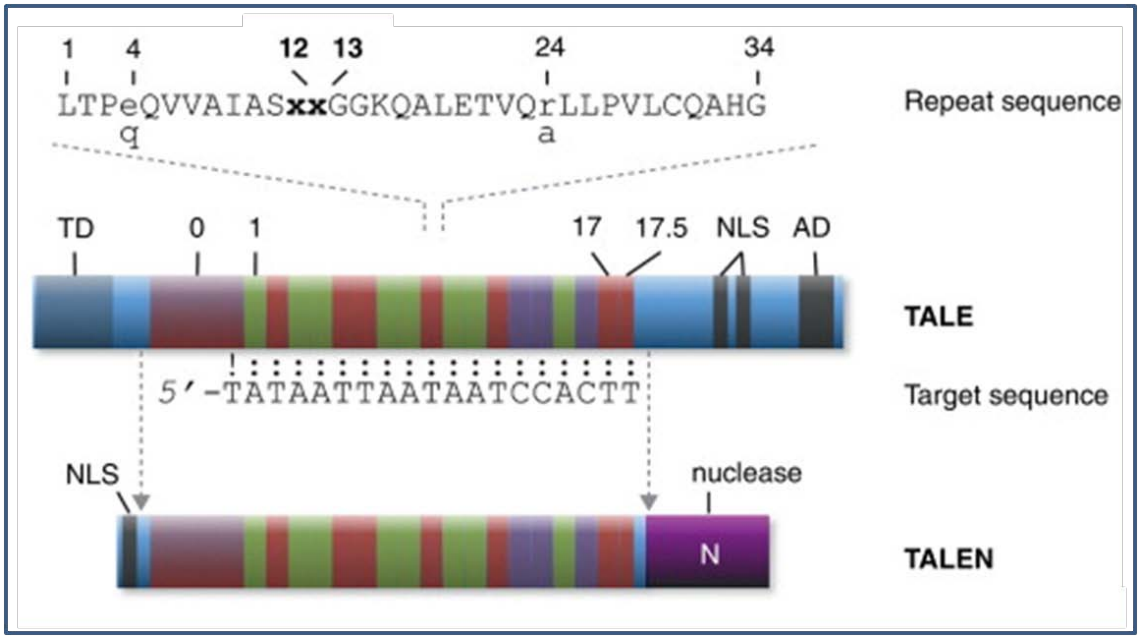



**Figure 1.11**: **Schematic diagram of TALEN structure and recognition site.** The basic TALE element includes: an N-terminal translocation domain (TD), the central repeat units (0–17.5) composed of ~34-35 multiple amino acids that recognises DNA target site, a C-terminal region comprising NLS, and a transcriptional activation domain (AD). The central TALE repeat units can be linked to a nuclease domain to produce a site-specific TALEN. The amino acids in positions 4 and 24 show some variability, while DNA specificity of the repeat unit is determined by the highly variable residues in positions 12 and 13 (repeat variable di-residue, RVD). Reprinted from Mussolino and Cathomen 2012.

## 1.5 Genome surgery by oligonucleotides

Alternatively, site-specific genome alterations at relative rates could be achieved in selected cell lines by designing short single stranded oligonucleotides (ssODN), chimeric RNA/DNA oligonucleotides and triple helix forming oligonucleotides (Igoucheva *et al*., 2004; Radecke *et al*., 2006). The ssODNs are able to introduce site

specific alterations and modify the target sequence at high frequencies (Olsen *et al*., 2009), hence, they represent an attractive option for gene repair. Some cell lines, like CHO and HEK 293T have shown site specific genetic alteration frequencies after transfection with ssODNs at about 0.03-0.5%, whereas primary mammalian cells demonstrated a very low frequency. Importantly, it has been shown that in mammalian cells the ssODNs achieve higher repair frequencies than double stranded DNA and the most favourable size for the ssODNs is between 20-80 bases (Olsen *et al*., 2005). The ssODNs basic design is a complementary sequence to the target locus which carries a central mismatch to the locus to be targeted. Despite the fact that the actual mechanism controlling the ssODNs action is not currently well understood, it has been shown the ssODNs integrate to the targeted locus during replication (Olsen *et al*., 2009).

## 1.6 Primary immunodeficiency

Primary immunodeficiencies (PID), also called inherited immune system disorders, are rare inherited disorders of the innate and acquired immune system. They result from more than 130 inherited mutations in genes required for production, differentiation and survival of specialised leukocytes like T or B lymphocytes, natural killer (NK) cells, neutrophils or antigen presenting cells (Aiuti and Roncarolo, 2009; Kohn, 2010; Notarangelo, 2010). Currently, more than 300 human genetic disorders are diagnosed as PIDs. Patients regulerly presented with a higher vulnerability to opportunistic infections or infection with uncommon organisms and may also develop autoimmunity or autoinflammatory disease and lymphoreticular malignancies (Rivat *et al*., 2012). PIDs are called "primary" because they are resulted from an inherited gene defect that affects the immune system. Secondary immune deficiencies are caused by acquired factors, such as chemotherapy or infection. PIDs, in general, are monogenic disorders following simple Mendelian inheritance while quite a few PIDs are polygenic. Monogenic diseases are those inherited disease that controlled by a single gene, while polygenic ones are more than one gene are involved. All the PIDs are rare (except the IgA deficiency) with incidence rate of approximately 1:10000 live births (Notarangelo, 2010). Usually, leukocytes are produced from the multipotent HSCs in bone marrow; so allogeneic hematopoietic stem cell transplantation (HSCT) is the preferred treatment for many PIDs (including SCID) (Kohn, 2010).

In spite of recent improvements in the allogeneic HSCT treatment, some patients are still suffering from lasting complications, and using of alternative donors is still associated with many difficulties. Alternatively, gene therapy approaches based on *ex vivo* HSCT by transferring functional genes into autologous HSCs using viral vectors, are emerged (Aiuti and Roncarolo, 2009). The first gene therapy treatment for PID was performed in 1991 for adenosine deaminase (ADA) SCID patient. Between 1991 and 2009, more than 90 patients have received gene therapy treatment for different PIDs (Aiuti and Roncarolo, 2009).

There are many possible advantages that make gene therapy treatment more attractive treatment:

- Reduction of graft-versus-host disease (GVHD) and minimising the related complications as gene therapy treatment based on gene correction and re-transplantation of the patient's own HSCs leading to the same clinical benefits as allogeneic HSCT, but without the immunologic complications (Kohn, 2010).
- Rapid access to treatment.
- HSCs abilities in renewal and proliferation could provide an important survival advantage for the gene-corrected cells leading to immunological repopulation and ultimately restoration of immune system (Qasim *et al*., 2009).
- HSCT can be easily manipulated *ex vivo,* and the *in vivo* vectors administration related problems can be bypassed (Qasim *et al*., 2009).

## 1.7 Hematopoietic Stem Cells (HSCs)

HSCs are rare cells of heterogeneous population capable of differentiating into blood lineages (Aguila *et al*., 2011). They are characterized by their capabilities in self-renewal, multi-lineage differentiation into different blood cells, and lifelong repopulation of haematopoiesis in lethally irradiated recipients (Trobridge and Kiem, 2010; Wognum *et al*., 2003). Thus, they are attracting a lot of interest for gene therapy applications and studies. Many life–threatening PIDs and acquired immune diseases can be treated successfully via gene transfer into HSCs by gene therapy (Mikkola *et al*., 2000). In healthy adults, different mature hematopoietic cells are normally substituted

by the proliferation and differentiation of multipotent HSCs mainly present in the bone marrow. It has been suggested that HSCs may also be found in other non-hematopoietic tissues like muscle, liver, vasculature and skin. Those HSCs to be used for gene therapy applications can be from different clinical sources, such as bone marrow, mobilized peripheral blood and umbilical cord blood (Kohn, 2010; van Hennik *et al*., 1998; Wagemaker *et al*., 1998). The relative ease of HSCs gene modification protocols in terms of simply harvesting of donor cells, transducing them with therapeutic vectors *ex vivo* and re-infusing them intravenously into an irradiated recipient; has made the hematopoietic gene therapy more applicable. This contrasts with gene therapy of solid organs which is more complex because of the *in vivo* or *ex vivo* trials to deliver genes with maximum transduction efficiency to large number of varied cell types within a complex tissue composition is very difficult (Trobridge and Kiem, 2010).

The features of human HSCs (like their cell-cycle characteristics, proliferative needs, purification techniques, engraftment kinetics, cytokines responsiveness, radiation sensitivity and homing abilities) are approximately similar to those of other higher mammals (Trobridge and Kiem, 2010; Wognum *et al*., 2003). Consequently, mouse HSCs are widely used as an attractive tools for hematopoietic studies establishing a solid knowledge of HSCs biology (Ema *et al*., 2006) . However; mouse HSCs are not useful for long-term engraftment assessment as they are short-lived animal model, and the variation between mouse and human HSC cluster of differentiation (CD) markers result in overestimation of gene therapy efficiency in mouse model (Trobridge and Kiem, 2010). The main differences between human and mouse HSCs can be summarised below:

- Division rate *in vivo*: mouse HSCs divide every 2.5 weeks; whereas human HSCs every 45 weeks (Shepherd *et al*., 2007), this is crucial when deciding the proper vectors that can transduce even the quiescent cells.

- Cell Cycle: about 75% of mouse HSCs are outside $G_0$ (Cheshier *et al*., 1999) whereas human HSCs cannot be subjected to cell cycle analysis because of the failure to obtain homogeneous cells (Trobridge and Kiem, 2010).

- Transduction efficiency: mouse HSCs can be transduced more easily than human ones (Trobridge and Kiem, 2010).

- Surface markers: there are different HSC CD markers; generally $CD34^+$ markers are used for human HSCs whereas $lineage^-$,$Sca\text{-}1^+$ and $c\text{-}kit^+$ are used for mice ones (Mikkola *et al*., 2000).

Theoretically, a single HSC can proliferate and expand to substitute the whole hematopoietic cells. However this concept is very difficult to be proved in practice mainly because a single HSC is not enough to re-generate the required quantities of hematopoietic cells to restore the entire haematopoietic system in patients or lethally irradiated animals. It has been reported that host (recipient) and donor HSCs can compete with each other in proliferation, engraftment and expansion in process known as "competitive repopulation"(Micklem *et al*., 1972). In mice, *ex vivo* HSCs protocols prefer to inject donor HSCs intravenously into the tail vain or subcutaneously. The process of migration, repopulation and localisation of these injected donor HSCs to the bone marrow is called "seeding efficiency" (Benveniste *et al*., 2003; Matsuzaki *et al*., 2004).

Importantly the HSCs and progenitors are not separate as there is a continuous proliferation in a hierarchy (Rosendaal *et al*., 1979). *In vivo*, adult HSCs are mostly quiescent and they possess very slow dividing kinetics, it is possible that maintenance of HSC quiescence and slow cell-cycle progression could be critically involved in sustaining a self-renewing HSC compartment for life (Qian *et al*., 2007). The quiescent nature of human HSCs has restricted gene transfer, since they are rather poor targets for common oncoretrovirus vectors which necessitate divided cells for integration (BouHamdan *et al*., 1999). This obstacle has been bypassed by using lentivirus vectors as their NLS elements enable the nuclear entry of PIC leading to transduce even the non-dividing cells (Engelman and Cherepanov, 2008); therefore lentivectors are recruited to transduce different HSC populations (Mikkola *et al*., 2000). The transduced vectors must not hinder HSCs features in renewal, differentiation or expansion; on the other hand vectors genome must be successfully sustained and transferred into HSCs progeny (Trobridge and Kiem, 2010).

## 1.8 SCID disorder

### 1.8.1 SCID description

Severe combined immunodeficiency (SCID) disorders, commonly called “bubble boy disease” (Kandavelou *et al*., 2005), are inherited fatal monogenic primary immunodeficiencies (Aiuti *et al*., 2007). SCID can be defined immunologically as a severe lack of specific immune response and clinically as a lethal disease in infants because of high susceptibility to opportunistic infections (Niehues *et al*., 2010). So far, there are at least 15 different genetic mutations causing different SCID disorders (Wiegant *et al*., 2010). Thus, SCID patients are a varied group of rare disorders characterized altogether by a severe decrease or lack of functional T cells (Alexander *et al*., 2007). Occasionally, there is a direct or indirect B cell dysfunction, and in some other cases NK cells will be affected as well. About 20% of SCID patients suffer from partial or complete lack of B and T lymphocytes. SCID disorders are classified into two categories according to the immunological phenotype, first is SCID with absence of T cells but presence of B cells ($T^-B^+$ SCID) and second with absence of both T and B cells ($T^-B^-$ SCID), (Notarangelo, 2010). The estimated general incidence rate is ranged between 1:50,000 and 1:100,000 live births; whilst the actual rate may be increased slightly if the early deaths are already counted.

Noteworthy, different features have been used to diagnose at least 11 diverse SCID conditions (Cavazzana-Calvo and Fischer, 2007). The most frequent SCID condition is the X-linked form (SCID-X1), which represent about 40-50% of all cases. SCID-X1 can be attributed to the defect in the gene encoding the interleukin-2 receptor γ chain (IL2RG). This molecule is a major subunit of the cytokine receptor complex for IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21 and is termed the common cytokine receptor γ chain (γc) (Gaspar *et al*., 2011a). SCID-X1 patients suffer from profound absence of cellular and humoral immunity, specifically lack T and NK cells with normal or high B cells numbers that are incapable of undergoing immunoglobulin maturation and antibody generation (Ginn *et al*., 2010). Because of profound growth of native HSCs and selective survival advantage that could be presented in corrected cells, SCID-X1 is a predominantly attractive target for gene therapy (Gaspar *et al*., 2011a).

The other common SCID disorder is ADA deficiency (ADA SCID) which represents about 10%-15% of all SCID forms. The absence of ADA enzyme leads to accumulation of high intracellular levels of toxic phosphorylate metabolites of adenosine and deoxyadenosine (ADA) substrates enhancing apoptosis of lymphoid precursors in the bone marrow and thymus (Notarangelo, 2010). As a result, abnormal T, B, and NK lymphocyte development and function is noticed in addition to other related systemic defects. Patients are routinely presented with severe opportunistic infections, failure to thrive, and a typical SCID profile. ADA-SCID could be treated by gene therapy as the regulation of ADA gene expression is relatively uncomplicated, and ADA-corrected cells have a survival advantage over mutated cells (Gaspar *et al*., 2011b).

Importantly, the first human patient with mutant DNA PKcs has been identified recently (van der Burg *et al*., 2009a). This DNA-PKcs deficiency was considered the first human *PRKDC* SCID patient. Phenotypically the patient presented with $T^-B^-$ SCID; and immunologically she had a severe lack in B and T cell numbers in peripheral blood and reduced numbers of B cell progenitors in the bone marrow. The patient cells lacked normal V(D)J recombination due to a deficient NHEJ repair mechanism (see section 1.8.2 below), (van der Burg *et al*., 2009b). The human *PRKDC* SCID patient (DNA-PKcs deficient) has two homozygous variations; the first is 3-nucleotide deletion causing the deletion of a glycine and the second is a missense mutation causing an amino acid substitution. The second mutation was confirmed to be the disease causing mutation (van der Burg *et al*., 2009b).

In the vast majority of SCID patients clinical findings are characterized by severe opportunistic infections (regularly in the respiratory tract), chronic diarrhoea accompanied by opportunistic microbes such as *Pneumocystis carinii* and *Aspergillus spp.* and intracellular organisms such as *Cytomegalovirus* can cause frequent infections and a failure to thrive. Unfortunately, without intensive medical care, death usually happens within the first year of the life (Cavazzana-Calvo and Fischer, 2007; van der Burg *et al*., 2009a). SCID is usually treated by bone marrow transplantation from human leukocyte antigen (HLA) matched sibling donor. Unfortunately, some patients do not have a proper donor leading to HLA mismatched transplant and eventually related with high risk of morbidity and mortality. Generally this treatment is associated

with insufficient immunological reconstitution, especially B cell function, and permanent requirement for immunoglobulin substitute therapy. Possible improved treatment has been provided by gene therapy promising a high survival rates and more comprehensive immunological reconstitution (Ginn *et al*., 2010).

SCID disorder can be cured by allogeneic HSCT; the first successful attempt was reported in 1968. Thus; by using allogeneic HSCT, hundreds of SCID patients have restored their healthy immune system around the world with about at least 80% chance of success when HLA-matched sibling donor is available and 70% in the case of availability of HLA-matched unrelated donor. However; the success of this approach of SCID treatment is limited because of different factors:

- It is highly dependable on the patient's age, older recipients usually related to higher failure rate.
- Immunological problems related to engraftment rejection
- Slower and occasionally incomplete immunological reconstitution
- Long-term decline of functional T cells.

Alternatively, gene therapy has emerged as a promising approach in SCID treatment exploiting the principle of genetic modification of autologous HSCT using retroviral vectors encoding for corrective genes. As of 2008, a total of 20 SCID-X 1 patients have been subjected to this treatment in London and Paris in two independent trails; 17 of them have effectively been cured. Out of these 17, 5 patients have developed leukaemia because of treatment-related insertional mutagenesis and unfortunately 1 of leukemic patiants has died (Cavazzana-Calvo and Fischer, 2007; Howe *et al*., 2008).

### 1.8.2 *PRKDC* SCID and NHEJ machinery

B and T cells can detect foreign antigens via the acquisition of specialised receptors, during lymphocyte maturation; these acquired receptors are immunoglobulin and T-cell receptors (TCR), respectively (Moshous *et al*., 2001). The profound immune disorder in SCID patients is a result of the failure to rearrange these receptors, which blocks lymphocyte maturation process causing lymphopenia and hypogammaglobulinemia (van der Burg *et al*., 2009a). In these receptors; the large diverse antigen-recognition areas are made of variable (V), diversity (D), and joining (J) gene segment, which

eventually go through rearrangement changes before their expression by a process called V(D)J recombination. V(D)J recombination is vital for both antigen recognition and development of lymphocytes (Niehues *et al*., 2010). Therefore; V(D)J recombination is considered a critical process for immune system maturation (Moshous *et al*., 2001).

At the borders of each of V, D and J segments there are recombination signal sequences (RSSs), where the V(D)J recombination is initiated by introduction of DSBs (van der Burg *et al*., 2009a), by the action of lymphoid-specific recombination activating gene (RAG) endonucleases 1 and 2 ( RAG is referred to RAG 1 and RAG 2 together), (Rooney *et al*., 2002). RAG 1 and 2 are strongly related to each other and are situated on human chromosome 11p13, producing 1043 and 527 amino acids respectively (Niehues *et al*., 2010). RAG induces DSBs at RSSs of the segment V, D or J generating two intermediates; a hairpin-sealed DNA coding ends on the chromosome and blunt signal ends removed from the chromosome (Rivera-Munoz *et al*., 2009). The next step is recognition and repair of the generated DSBs by the general NHEJ machinery. The mammalian NHEJ repair machinery is composed of at least 7 known proteins: Ku70, Ku80, DNA-PKcs, Artemis (Art), XRCC4, Cernunnos/XLF, and DNA ligase IV (lig 4), (Rivera-Munoz *et al*., 2009; Rooney *et al*., 2002). DNA-PK is a complex protein, containing heterodimeric Ku protein (Ku70 and Ku80 subunits of 70 and 86 kDa, respectively) and DNA-PKcs (Beamish *et al*., 2000). The Ku70 and Ku80 subunits can bind to DNA DSB ends, nicks and DNA hairpins; whereas DNA-PKcs is completely dependent, for activity, on binding to DNA DSB including broken ends, nicks and single stranded gaps (Kirchgessner *et al*., 1995). Ku proteins can act as component of DNA-PKcs (Blunt *et al*., 1995).

The earlier recognition of DSBs is achieved by Ku70/Ku80 heterodimers, where the DNA-PKcs is recruited producing DNA end synapsis, to guarantee protection from exonuclease activities (van der Burg *et al*., 2009a). The final step of V(D)J recombination is to re-ligate the DNA DSBs of the two broken ends of the chromosome by the other NHEJ proteins. Thus, any mutations in the genes coding for NHEJ proteins will result in different SCID disorders in humans. Obviously, in mammalian cells the NHEJ repair machinery acts as a key pathway in DSBs repair. For that reason SCID patient are hypersensitive to any DSBs inducer regents like ionizing radiations.

Likewise SCID mice suffer from severe dysfunction in B and T cells differentiation (Rivera-Munoz *et al*., 2009).

### 1.8.3 Spontaneous mutations in mouse *Prkdc* gene

The spontaneous mutations in the *Prkdc* gene were identified first in *scid* mice; nevertheless *scid* mice are not the only known *scid* animal; as these mutation are found in dogs (Trobridge and Kiem, 2010), horses of Arabian breed and very recently in human as well (van der Burg *et al*., 2009b). The *scid* mice, are lacking recombination process needed for both DSB repair and V(D)J recombination (Kirchgessner *et al*., 1995). Phenotypically, they suffer from cellular hypersensitivity to ionizing radiation and severe deficiency in B and T cells. Murine *scid* disorder is attributed to a spontaneous point mutation in the *Prkdc* gene (Al-Wahiby *et al*., 2005), a gene coding for DNA-PKcs comprising of 86 exons extending across about 193 kb and is located on chromosome 16 (Gilley *et al*., 2001). It is worth mentioning that, the point mutation of the *scid* mouse was corrected previously using complementation of *Prkdc* gene. A full copy of the gene cDNA was cloned into the yeast artificial chromosome (YAC) and used to transfect immortalised mouse *scid* cell line leading to correct *Prkdc* and restore DNA-PKcs function (Priestley *et al*., 1998).

The DNA-PKcs is a serine/ threonine protein kinase activity which belongs to the phosphoinositol 3-kinase family (Niedernhofer, 2008), and composed of ~465 kDa (Hande *et al*., 1999). In the *scid* mice, defective *Prkdc* gene switches the Tyr-4046 codon into a stop codon resulting in truncation of 83 amino acids, losing the highly conserved C-terminal region (Al-Wahiby *et al*., 2005); *scid* mice have a profound impairment V(D)J recombination. The function of DNA-PKcs is initiated by the binding of Ku70/Ku80 heterodimers to the DNA ends of the DSB leading to form the DNA PK complex (Ku70, Ku80 & DNA-PKcs). This complex will acquire the activity to phosphorylate many proteins involved in NHEJ repair including itself. This self-phosphorylation process is vital in NHEJ repair mechanism (van der Burg *et al*., 2009c).

The *scid* mice first reported in 1983 by Bosma *et al* (Bosma *et al*., 1983) , are characterised by their severe lack of T and B cells but with functional NK cell encouraging many researchers to use *scid* mice as a model for different studies.

Interestingly, *scid* mice are also used to develop tissues and cell lines, which show similar features to their original *scid* mice in terms of hypersensitivity to gamma irradiation and a fault in the repair of DSBs. However, it has been reported that *scid* mice are not absolute mutants as many of them show low DNA-PKcs activity. Those "leaky" *scid* mice represent about 2-25% of young adult of the *scid* population and nearly all old *scid* mice. In the laboratories, *scid* mice can be housed under conventional facilities for only 5-6 months, as like their human patient counterparts, they fall victim to opportunistic infections. However, if they housed under pathogen-free environment these animals can survive for up to 2 years (Leblond *et al*., 1997).

## 1.9 Aims of the project

HSCs are an attractive target for gene therapy, as they can both self-renew and differentiate into all blood lineages, and the possibility to manipulate them *ex vivo*. Gene transfer into HSCs can potentially provide a cure for many inherited and acquired diseases of the hematopoietic and immune systems. Thus, corrected HSCs could support hematopoiesis throughout the lifetime. So far, conventional retroviral vectors were used to treat the hematopoietic disorders by gene therapy and the success of these attempts has been limited by insertional mutagenesis mainly due to the natural tendency of these retroviruses to integrate into the host genome.

This study could be considered as an attempt to improve the accuracy of gene therapy treatment. It will use gene repair to correct classical *scid* mouse which recently has become a model of human *PRKDC* SCID (DNA-PKcs deficiency). A combined technology of ZFNs and lentiviral vectors will be applied to correct *scid* mouse models by using HR mediating gene repair. The dysfunctional *Prkdc* gene in the HSCs will be targeted at a very close region to the *scid* mouse point mutation by ZFNs to induce site-specific DSBs. This aim can be achieved by *ex vivo* transduction of these cells with lentiviral vectors carrying two different transgenes; the first is ZFN to induce site-specific DSBs, which will provoke the HR repair machinery and the second is a functional homologous sequence of the wild type *Prkdc* gene, which will act as a corrective template for the HR machinery.

A short time after the *ex vivo* transduction, *scid* HSCs will be transplanted back into the *scid* mouse recipients. When the corrective template replaces the faulty *Prkdc* sequence and expression restored in the transplanted cells, the corrected cells should display significant growth as they possess a selective survival advantage. In order to reach this aim and to generate even a sustained effect, the corrected HSC progenitors must have a self-renewal capacity and at the same time repopulate the T- and B-cells, i.e. keep their multipotency. In fact, the multipotency and selective advantage of corrected HSC progenitors are hypothesized to be the major factors counting for the success in this project. By transplantation of the corrected cells into recipient *scid* mice and by relying on their abilities to repopulate with a selective survival advantage, the *scid* mouse is hoped to be corrected perfectly.

The current project intends, for the first time, to show that gene repair could be considered as an efficient therapeutic strategy for the treatment of PIDs in haematopoietic animal model. Such therapeutic gene manipulation would be the ideal form of gene therapy for PIDs and many other genetic disorders because the repaired gene would be expressed under endogenous regulation, bypassing concerns about expression level, timing or stability.

The feasibility of this system will be tested and optimised first *in vitro* using *scid* fibroblast cell line. Two corrective donor templates will be used both carrying homologous sequences of *Prkdc* around the *scid* point mutation. Additionally, the first codes for positive selective marker to be used for *in vitro* selection and the second encodes a unique site for the enzyme *Bsa*WI to be used as a diagnostic site for *ex vivo* gene repair. The following steps will be used in this project to achieve the final aim:

1. Immortalisation of murine *scid* fibroblasts to avoid the extensive senescence noticed in these cells.
2. Incorporation of corrective templates into IDLVs.
3. Incorporation of ZFN gene constructs into IPLVs and IDLVs.
4. Testing and optimising the efficiency of ZFNs to induce site-specific DSBs in mouse *scid* fibroblasts.
5. First gene repair attempt *in vitro* using immortalized *scid* fibroblasts cell line and corrective constructs carrying diagnostic markers.
6. Analysis of ZFN cutting efficiency in mouse HSCs.
7. Main gene repair experiment *ex vivo* using *scid* HSCs and corrective constructs carrying a silent diagnostic restriction site.
8. Transplantation of repaired haematopoietic stem cells will be transplanted into *scid* mouse recipients.
9. Finally, engraftment and potential therapeutic effect in recipient mice will be analysed.

# Chapter Two

# Materials and Methods

# 2 Materials and Methods

## 2.1 Materials

All solutions, buffers and media were prepared with double deionised water ($ddH_2O$). Stock solutions were sterilised by autoclaving or passing through a 0.22 µm filter and then stored at room temperature (RT), unless otherwise stated.

### 2.1.1 Common Materials

Materials/reagents common to more than one protocol are listed below in alphabetical order:

#### 2.1.1.1 General Laboratory Chemicals

| Items | Supplier |
|---|---|
| 2-mercaptoethanol | Sigma-Aldrich, UK |
| AccuGel acrylamide solution (40 % w/v) | National diagnostic, USA |
| Acrylamide solution (30 %) | Merck, Germany |
| Agar, Noble | Sigma-Aldrich, UK |
| Agarose molecular grade | Bioline, UK |
| Albumin, from bovine serum | Sigma-Aldrich, UK |
| Alkaline protease | Promega, USA |
| Ammonium acetate | Sigma-Aldrich, UK |
| Ammonium persulfate | Sigma-Aldrich, UK |
| Ampicillin | Sigma-Aldrich, UK |
| Bacto-Trytone/Peptone | DIFCO, USA |
| Barium chloride | Sigma-Aldrich, UK |
| Boric acid | Sigma-Aldrich, UK |
| Bromophenol blue | Sigma-Aldrich, UK |
| Calcium chloride anhydrous | Sigma-Aldrich, UK |
| Crystal violet | Sigma-Aldrich, UK |
| DAPI | Sigma-Aldrich, UK |
| dN6 | Sigma-Aldrich, UK |
| Dried skimmed milk powder | Merck, Germany |
| DTT | Sigma-Aldrich, UK |
| EDTA | Sigma-Aldrich, UK |
| Ethanol | British Drug House (BDH) |
| Ethidium bromide | Invitrogen, USA |
| Ficoll type 400 | Sigma-Aldrich, UK |

| | |
|---|---|
| Formaldehyde | Sigma-Aldrich, UK |
| G418 | Sigma-Aldrich, UK |
| Gelatine | Sigma-Aldrich, UK |
| GeneScreen Plus membrane | Perkin Elmer, USA |
| Glacial acetic acid | Sigma-Aldrich, UK |
| Glycerol | Sigma-Aldrich, UK |
| Glycine | Sigma-Aldrich, UK |
| HEPES | Sigma-Aldrich, UK |
| Hybond™ ECL™ Nitrocellulose membrane | Amersham Biosciences |
| Hydrochloric acid | BDH, UK |
| IGEPAL | Sigma-Aldrich, UK |
| Kanamycin | Sigma-Aldrich, UK |
| Kodak Biomax light film | Kodak, USA |
| LB broth medium | Sigma-Aldrich, UK |
| Magnesium acetate | Sigma-Aldrich, UK |
| Magnesium chloride | Sigma-Aldrich, UK |
| Magnesium sulphate | Sigma-Aldrich, UK |
| Manganese chloride | Sigma-Aldrich, UK |
| Microspin S-300 column | GH Healthcare |
| MOPS | Sigma-Aldrich, UK |
| N2 supplement | Invitrogen, USA |
| Orange G | Sigma-Aldrich, UK |
| Paraformaldehyde | Sigma-Aldrich, UK |
| Phenol | Fisher Scientific |
| Phosphoric acid | BDH, UK |
| Polybrene | Sigma-Aldrich, UK |
| Potassium acetate | Sigma-Aldrich, UK |
| Potassium chloride | Sigma-Aldrich, UK |
| Potassium hydroxide. | Sigma-Aldrich, UK |
| Puromycin | Sigma-Aldrich, UK |
| Rubidium chloride | Sigma-Aldrich, UK |
| Sodium acetate | Sigma-Aldrich, UK |
| Sodium azide | Sigma-Aldrich, UK |
| Sodium bicarbonate | Sigma-Aldrich, UK |
| Sodium chloride | Sigma-Aldrich, UK |
| Sodium deoxycholate | Sigma-Aldrich, UK |
| Sodium dodecyl sulphate (SDS) | BDH, UK |
| Sodium fluoride | Sigma-Aldrich, UK |
| Sodium hydroxide | Sigma-Aldrich, UK |
| Sucrose | Sigma-Aldrich, UK |
| TEMED | Sigma-Aldrich, UK |
| Tris base | Sigma-Aldrich, UK |
| Tris hydrochloride | Sigma-Aldrich, UK |

| | |
|---|---|
| Trisodium citrate dihydrate | Sigma-Aldrich, UK |
| Triton X-100 | Sigma-Aldrich, UK |
| Tween 20 | Sigma-Aldrich, UK |

#### 2.1.1.2 Biological Kits

| Kit | Supplier |
|---|---|
| Accuprime™ Taq DNA high fidelity polymerase kit | Invitrogen, USA |
| Bradford method for protein detection | Bio-Rad, USA |
| Complete mini tablets protease inhibitor cocktail | Roche, Germany |
| Copmbeads™ anti-mouse Ig, k/negative control (FBS) compensation particles set | BD Biosciences, USA |
| DNA endonuclease Surveyor® mutation detection kit | Transgenomic, USA |
| DNeasy® extraction kit | Qiagen, Germany |
| ECL™ Western blotting detection reagents | Amersham Pharmacia, UK |
| Endotoxin-free Maxiprep kit , tip-500 | Qiagen, Germany |
| Endotoxin-free Megaprep kit, tip-2500 | Qiagen, Germany |
| Endotoxin-free Miniprep kit, spin columns | Qiagen, Germany |
| Lineage cell depletion kit | Miltenyi MACS, Germany |
| Qiagen PCR purification kit | Qiagen, Germany |
| Qiaquick® gel extraction kit | Qiagen, Germany |
| RBC lysis buffer | Biolegend, USA |
| SignaTECT® DNA-dependent protein kinase | Promega, USA |

#### 2.1.1.3 Markers and Restriction Enzymes

| Item | Supplier |
|---|---|
| 1 Kb ladder | Promega, USA or<br>New England Biolabs (NEB), UK |
| DNA modifying enzymes | NEB, UK |
| HyperLadder™ I | Bioline, UK |
| HyperLadder™ V | Bioline, UK |
| Loading DNA Buffer (6 X) | NEB, UK |
| Novex® Sharp Pre-stained Protein Standard | Life Technologies, UK |
| Restriction enzymes and buffers | NEB, UK |

### 2.1.1.4 PCR and Southern blotting reagents

| Item | Supplier |
|---|---|
| [α-$^{32}$P] dATP | Perkin Elmer, USA |
| [γ-$^{32}$P] ATP | Perkin Elmer, USA |
| ABI universal master mix | Bioline, UK |
| GoTaq DNA polymerase | Promega, USA |
| Longamp Taq DNA polymerase | NEB, UK |
| SYBR green | Bioline, UK |
| Taqman master mix | Roch, Germany |
| β-actin probe | Applied Biosystems, USA |

### 2.1.1.5 Tissue culture

| Item | Supplier |
|---|---|
| 0.22 µm-pore-size filter | Nalgene, USA |
| 30 µm nylon mesh filters | Millipore, USA |
| DMSO | Sigma-Aldrich, UK |
| Dulbecco's modified Eagle medium DMEM | PAA, Austria |
| Dulbecco's phosphate buffer saline (PBS) | PAA, Austria |
| Foetal bovine serum (FBS) | PAA, Austria |
| High-speed polyallomer centrifuge tube | Beckman Coulter, USA |
| Human IL-6 | PeproTech, USA |
| Melphalan | Sigma-Aldrich, UK |
| MTT | Sigma-Aldrich, UK |
| Murine Flt-3 | PeproTech, USA |
| Murine stem cell factor (SCF) | PeproTech, USA |
| Penicillin / streptomycin | PAA, Austria |
| StemSpan® medium | Stemcell Technologies, Canada |
| Trypsin/EDTA | PAA, Austria |

#### 2.1.1.6 Antibodies

| Item | Supplier |
|---|---|
| Alexa fluor 680Goat anti-rabbit secondary antibody (cat ≠A-21076) | Invirogen, USA |
| Anti-Human/Mouse CD45R (B220) PE (cat ≠12-0452-81) | eBioscience, USA |
| Anti-Mouse CD11b APC (cat ≠17-0112) | eBioscience, USA |
| Anti-Mouse CD3e PerCP-Cy5.5 (cat ≠45-0031-82) | eBioscience, USA |
| Anti-Mouse CD4 PE (cat ≠12-0041) | eBioscience, USA |
| Anti-Mouse CD8a APC (cat ≠17-0081) | eBioscience, USA |
| Armenian Hamster IgG Isotype Control PerCP-Cy5.5 (cat ≠45-4888) | eBioscience, USA |
| HRP-conjugated polyclonal goat anti-mouse secondary antibody (cat ≠P044701-2) | Dako, UK |
| IRDye 800CWgoat anti-mouse secondary antibody (cat ≠926-32210) | Li-cor GmbH,Germany |
| Mouse anti-FLAG M2 monoclonal primary antibody (cat ≠200472) | Stratagene, UK |
| Mouse monoclonal anti α-tubulin primary antibody (cat ≠T9026) | Sigma Aldrich,UK |
| Rabbit polyclonal to α-tubulin primary antibody (cat ≠ab18251) | Abcam, UK |
| Rat IgG2a K Isotype Control PE (cat ≠12-4321) | eBioscience, USA |
| Rat IgG2b K Isotype Control APC (cat ≠17-4031) | eBioscience, USA |

### 2.1.2 Buffers for agarose gel electrophoresis

TAE (50 X stock) buffer was prepared by dissolving 242 g of Tris base (2-amino-2-hydroxymethyl-propane-1, 3-diol), 57.1 ml of glacial acetic acid and 100 ml of 0.5 M EDTA (pH 8.0) in 600 ml of $ddH_2O$. The pH was adjusted to 7.6, and then the final volume was made up to 1000 ml. A working solution was prepared by diluting to 1 X with $ddH_2O$. TBE (10 X stock) buffer was prepared by dissolving 108 g Tris base, 55 g of boric acid, and 40 ml of 0.5 M EDTA (pH 8.0) in 600 ml of $ddH_2O$. The pH was adjusted to 8.3, and then the final volume was made up to 1000 ml. A working solution was prepared by diluting to 1 X with $ddH_2O$.

## 2.2 Methods

### 2.2.1 Bacterial Manipulation

#### 2.2.1.1 Growth and Maintenance of *E. coli*

*Escherichia coli* (*E. coli*), Top 10 strain, were grown in Luria-Bertani (LB) (Gellhaus *et al*., 2010) broth medium at 37 ºC with agitation at 250 rpm or streaked out on solid LB plates containing 1.5 % bacto agar. Depending on the antibiotic resistance conferred by transforming plasmids, both LB broth and agar media were supplemented with either ampicillin (100 µg/ml) or kanamycin (50 µg/ml) as final concentrations. Stock solutions of ampicillin (100 mg/ml) and kanamycin (10 mg/ml) were made up using sterile dd$H_2O$, filtered, aliquoted and stored at -20 ºC. LB agar media were sterilised by autoclaving, then left to cool down to ~45 ºC before addition of antibiotic and pouring of plates.

For liquid cultures individual *E. coli* colonies were retrieved from fresh culture plates using a sterile inoculation loop and dispersed in 1-2 ml of LB broth. For plated cultures an inoculation loop was used to streak out bacteria across the surface of the agar plate. Culture plates were then incubated at 37°C for 16- 18 h. Plates were then sealed by wrapping in parafilm and stored at 4 °C.

#### 2.2.1.2 Glycerol stocks

For long-term storage at - 80 °C, bacterial cultures were supplemented with glycerol to 15 %.

#### *2.2.1.3* Transformation of *E. coli*

##### *2.2.1.3.1* Preparation of chemically competent *E. coli*

Chemically competent *E. coli* were prepared by streaking out LB agar plates with loops inoculated with *E. coli* from glycerol stocks and incubating overnight at 37 °C. Then, a mini-culture was prepared by picking up a single colony to inoculate 5 ml of LB broth and overnight incubation with agitation. The 5 ml mini-culture was mixed with 10 ml LB broth and incubated at 37 °C with agitation for ~ 1.5 h until the culture reached an optical density of ~0.5 at 600 nm wavelength. The cells were collected by centrifugation at 1000 X *g* for 5 min at 4 °C. The pellets were then resuspended in 15 ml of cold TFB1

buffer (see below for recipe) and incubated on ice for 15 min. Afterwards, the mixture was centrifuged at 1000 X *g* for 5 min at 4 °C. The pellets were then resuspended in 5 ml of ice cold TFB2 buffer (see below for recipe) and incubated 15 min on ice. Finally, 100 µl aliquots in microfuge tubes were placed in liquid nitrogen for a few seconds and then stored at -80 °C.

Buffer TFB1 was: 100 mM Rubidium chloride, 50 mM manganese chloride, 30 mM potassium acetate, 10 mM Calcium chloride, 15 % (v/v) glycerol, pH adjusted to 5.8 with potassium hydroxide. TFB2 buffer was: 10 mM MOPS (3-N-morpholino-propanesulfonic acid), 10 mM Rubidium chloride, 75 mM Calcium chloride, 15 % (v/v) glycerol, and pH adjusted to 6.8 with potassium hydroxide. Both TFB1 and 2 solutions were sterilised by filtration through 0.2 µm filters.

#### 2.2.1.3.2 Heat-shock bacterial transformation

Heat-shock mediated plasmid DNA transformation was performed by initially thawing chemically competent *E. coli* on ice. Plasmid ligation mixture was added to 100 µl of competent cell suspension and incubated on ice for 30 min. The cells were then heat-shocked at 42 °C for 45 s in a water bath and then immediately placed back on ice for 10 min. The cells were then diluted with 1 ml of plain LB medium and incubated at 37 °C for 60 min at 250 rpm using a thermo-block shaker, prior to spreading onto agar plates containing an appropriate selective antibiotic.

#### 2.2.1.3.3 Screening of bacterial transformants

About 100-200 µl of transformed bacterial suspension was spread onto LB agar plates, supplemented with the selective antibiotic matching the antibiotic resistance of the corresponding plasmid backbone. Colonies were allowed to grow overnight at 37 °C. An individual, isolated colony was picked from the plate using a sterile pipette tip, and used to inoculate 5 ml of LB/antibiotic media. The culture was then incubated overnight at 37 °C. DNA was subsequently isolated from 1.5 ml of culture using a Qiagen Miniprep kit, and characterised by restriction analysis followed by gel electrophoresis. For cloning purposes, batches of 20 colonies were typically tested. Clones displaying appropriate restriction maps were inoculated to larger culture volumes of LB/antibiotic for overnight culture at 37 °C with agitation. Subsequently Endotoxin-free Maxiprep or

Megaprep-scale plasmid purification was carried out according to the kit manufacturer's instructions (see section 2.1.1.2).

#### 2.2.1.4 Plasmid DNA preparation

##### 2.2.1.4.1 Small-scale purification of plasmid DNA

Small-scale plasmid DNA purification was achieved by alkaline lysis using Qiagen Miniprep kits as per manufacturer's instructions from 5 ml of overnight culture.

##### 2.2.1.4.2 Large-scale purification of plasmid DNA

For large-scale plasmid DNA preparation, alkaline lyses were done using Endotoxin free Maxi- and Megaprep kits (Qiagen) according to manufacturer's instructions. A 10 ml LB/antibiotic starter culture was set up from residual cultures previously checked by Miniprep purification and restriction analysis. Starter cultures were grown at 37 °C for 8 h. These were then diluted 1:1000 in LB/antibiotic. Maxi-cultures were prepared in 100 ml, and mega-cultures in 500 ml of LB media. Briefly, the preparation was conducted by binding of DNA to the anion-exchange resin columns under low-salt and pH conditions with an integrated endotoxin removal step. A medium-salt wash used to remove protein and other low molecular weight impurities. Finally, plasmid DNA was eluted from the column using high-salt buffer, and was subsequently desalted and concentrated by precipitation with isopropanol. The DNA was allowed to dissolve in 0.5-1 ml sterile ddH2O or TE buffer at room temperature and then stored at -20 °C.

##### 2.2.1.4.3 Measurement of DNA concentration

Plasmid DNA concentration was calculated by measuring the absorbance of light with a wavelength of 260 nm (A 260) using a NanoDrop ND-1000 spectrophotometer with a 0.2 mm path length. At this wavelength 50 µg/ml of double-stranded DNA has an absorbance of 1. Samples were run on an agarose gel against DNA molecular weight markers to double-check the concentration and DNA integrity.

### 2.2.2 Gel electrophoresis

#### 2.2.2.1 Agarose gel electrophoresis

Mixtures of digested DNA/loading buffer were loaded into the wells of an agarose gel prepared in TAE buffer (see section 2.1.2 above). An appropriate DNA molecular weight marker and undigested samples of plasmid DNA were run alongside the digested samples. All samples were then electrophoresed in 1 X TAE electrophoresis buffer. Ethidium bromide (EtBr) was added to the gel at a final concentration of 0.5 µg/ml prior to casting, in order to facilitate visualisation of the DNA fragments under UV light. Gel images were captured electronically using a Bio-Rad Gel Doc 2000 system.

Two types of electrophoresis gels were used: analytical gels (0.7 % agarose gel, run at 50-60 V) to check the presence and concentration of digested DNA bands, and preparative gels (0.8-1.5% agarose gel, run at ~32 V) to separate required DNA fragments during the process of cloning. Prior to ligation reactions, comparison of visualised DNA bands with bands of known concentration within HyperLadder® molecular weight controls was used to estimate the relative amounts of vector and insert fragments.

#### 2.2.2.2 Polyacrylamide gel electrophoresis (PAGE)

PAGE was used to analyse DNA fragments of less than 1000 bp in size. For double stranded DNA fragments of 50-300 bp, a 10 % acrylamide gel was prepared using a Minigel™ apparatus (Bio-Rad, UK). 2.5 ml of AccuaGel® 29:1 (National Diagnostic, USA), 7.4 ml of TBE 1 X, 100 µl of freshly prepared 10 % w/v ammonium persulfate (APS), and 5 µl Tetramethylethylenediamine (TEMED) were mixed and used for casting of vertical gels. The loading buffer was prepared with 15 % (w/v) Ficoll type 400 (Sigma Aldrich) and 0.25 % (w/v) orange G (Sigma Aldrich) kept frozen at -20 °C. Usually, the gels were run at 100 V, and then they were stained with 5 µl EtBr of 10 mg/ml in 100 ml 1 X TBE buffer for 10 min and then quickly washed several times with $ddH_2O$. Images were captured with UV illumination using a Bio-Rad Gel Doc 2000 system.

### 2.2.3 Molecular cloning methods

#### 2.2.3.1 Digestion of DNA by restriction endonuclease

Restriction enzyme digestion was used to produce compatible DNA ends for subcloning to plasmid vectors and to verify the resulting DNA vectors. For production of plasmids and inserts for subcloning, 1-10 µg of plasmid DNA was digested with appropriate restriction enzyme and fragments were isolated by agarose gel electrophoresis (see section 2.2.2.1 above) and purified from gel by purification with an appropriate kit (Qiaquick, Qiagen, Germany). To verify the quality of DNA plasmid preparations, approximately (~0.5 µg) of plasmid DNA was digested.

DNA was either linearised with a restriction enzyme corresponding to a unique site in the construct, or subjected to multiple enzymatic digestion. In the latter case, if the enzymes had compatible reaction conditions they were added to the same reaction; otherwise the digestions were performed separately for each enzyme with a purification (PCR purification kit, Qiagen-Germany) step in between. Digests were carried out using the endonuclease manufacturer's (NEB or Promega) recommended conditions and buffers in a 15 µl total reaction volume. Prior to electrophoresis, a sample of digested plasmid DNA (~200ng) was diluted with $H_2O$ and mixed with the appropriate volume of 6x loading buffer. Vector NTI software (Invitrogen, UK) was routinely used to predict restriction patterns.

#### 2.2.3.2 Alkaline phosphatase treatment

Following restriction enzyme digestion, to minimise the possibility of vector self-ligation, the 5'phosphate residue groups were removed from both termini of the linear, double-stranded recipient plasmid vector DNA. This was achieved using Antarctic™ phosphatase enzyme (NEB, UK), which catalyzes the removal of 5´ phosphate groups from DNA. As phosphatase-treated fragments lack the 5´ phosphoryl termini required by ligase, they cannot self-ligate. 1 µl of Antarctic™ phosphatase was added to 20-50 µl of digested DNA and the mixture was incubated at 37 °C for 15 min and then stopped by heat inactivation at 65 °C for 5 min. Later, phosphatase-treated fragments were electrophoresed onto preparative agarose gel and the DNA was purified using Qiaquick gel extraction kit, Qiagen-Germany.

#### 2.2.3.3 5’ overhang removal

To convert 5’ overhangs to blunt-end termini, Klenow treatment was used. Digested DNA was generally purified (Qiagen, Germany), adjusted to DNA polymerase buffer and 40 µM dNTPs, and 1 unit of Klenow polymerase (NEB, UK) was added per 1 µg of DNA. Alternatively, whenever the restriction digest buffer was compatible with the function of the Klenow polymerase, the enzyme and dNTPs were added directly to the restriction digest mixture. The reaction mixture was incubated at RT for 15 min and then stopped by heat inactivation at 75 °C for 20 min.

#### 2.2.3.4 3’ overhang treatment

Those DNA fragments with 3’overhangs were treated with T4 DNA polymerase for end-filling. A1 µl of T4 DNA polymerase (NEB, UK) together with 40 µM of dNTPs was added to the DNA mixture and incubated at 12 °C for 15 min and then stopped by heat inactivation at 75 °C for 20 min.

#### 2.2.3.5 Ligation-related preparation of cohesive and blunt termini

DNAs were digested with the appropriate restriction enzyme(s) to produce compatible ends for cloning. DNA inserts can be re-joined by ligase enzyme to a cohesive terminus of a linear vector DNA fragment previously cut with a compatible restriction enzyme or to a blunt-ended plasmid DNA fragment, frequently resulting from Klenow treatment. The products of the ligation reaction are introduced to competent *E coli* via transformation. The number of "correct" clones can be maximised by using insert/ vector molar ratios of 3:1 for cohesive termini, and 5:1 for blunt termini ligation. The ligation reactions were usually set up by using 100 ng of the vector DNA as a benchmark.

#### 2.2.3.6 DNA retrieval from gels

The plasmid fragments (vector/insert) were electrophoresed in EtBr-containing gels of appropriate density for the purpose of band size fractionation. Larger fragments (> 1000

bp) were electrophoresed into less dense 0.8 % (w/v) agarose gels, while for smaller fragments 1-1.5 % agarose gels were used. After electrophoresis, agarose bands containing the fragments of choice were retrieved by excision from the gel under UV radiation (EtBr-visualisation) using a sterile scalpel blade. DNA purification was done typically using a Qiaquick gel extraction kit (Qiagen, Germany). Finally the fragments were routinely dissolved in 30 µl of sterile TE and kept frozen at -20 °C.

#### 2.2.3.7 DNA ligation

T4 DNA ligase catalyzes the formation of a phosphodiester bond between a 5' phosphate group and a 3’ hydroxyl terminus in duplex DNA. Purified DNA insert fragments were ligated into linearised/dephosphorylated vector fragments in a 10 µl reaction volume using bacteriophage T4 DNA ligase (NEB, UK) in 1 X T4 ligation buffer at 16 °C for 16 h. The ligation reactions were set up using 100 ng of the vector DNA as a benchmark .The molar ratio of (insert: vector) fragments was 3:1 for cohesive-end ligations; whereas for blunt-end ligations it was 5:1. For conversion of molar ratios to mass ratios, the calculations were done by using the formula:

Molar ratio insert/vector = [Vector (ng) x insert size (kb)] / [insert (ng) x vector size (kb)]

### 2.2.4 Tissue culture methods

#### 2.2.4.1 Cell lines

Human and murine derived cell lines used for the purposes of this thesis are listed below. Further details of the features of the cell lines can be found at the European Collection of Animal Cell Cultures (ECACC, http://www.ecacc.org.uk) and the American Type Culture Collection (ATCC, http://www.lgcstandards-atcc.com).

**HEK-293T:** an adherent epithelial human embryonic kidney cell line, characterised by the presence of integrated copy of the human adenovirus serotype 5, E1a gene and constitutive expression of SV40 large T antigen (Graham *et al*., 1977). HEK 293T cells were used in this work to produce lentiviral and retroviral vectors.

**HeLa:** an adherent cervical epithelial cell line derived from a human adenocarcinoma, first described by Scherer (Scherer, 1954). Its morphological aspects are described in (Macville *et al*., 1999). HeLa cells were used for titration of lentiviral vectors.

**BALB/c 3T3 fibroblast**: 3T3 cells come from a cell line established by Todaro and Green in 1962. The 3T3 cell line was originated from primary Swiss mouse embryonic fibroblasts. The '3T3' nomenclature refers to growth condition of "**3**-day **t**ransfer, inoculum **3** x $10^5$ cells" which is required to maintain the cells continuously. The spontaneously immortalized cells with stable growth rate were established after 20-30 generations in culture, and then named '3T3' cells (Todaro and Green, 1963). 3T3 cells are adherent murine embryonic fibroblasts and they possess the ability to divide indefinitely, but are highly sensitive to the post-confluence inhibition of cell division. The cells grow in a monolayer, ceasing division when they cover the dish surface (Aaronson and Todaro, 1968). They were kindly provided by John Thacker (Medical Research Council, Radiation & Genome Stability Unit, Oxon- UK).

**BALB/c 3T3 *scid* fibroblast cell line:** Adherent murine embryonic fibroblasts established from *scid* BALB/c mouse whole embryo cultures with similar characteristics to murine BALB/c 3T3 fibroblasts. These cells were also provided by John Thacker (Medical Research Council, Radiation & Genome Stability Unit, Oxon-UK).

**BALB/c 3T3 *scid mTert* fibroblast cell line:** Adherent murine embryonic fibroblasts established from murine BALB/c 3T3 *scid* fibroblast cell line by immortalisation (please see section 2.2.8 below).

**Murine hematopoietic stem cells** (**wt and *scid*)**: non- adherent primary cell culture, prepared as described in section 2.2.4.2 below.

### 2.2.4.2 Cell line maintenance in culture

HEK-293T and HeLa cell lines were normally maintained in Dulbecco's modified Eagle medium (DMEM) with stable glutamine and 4.5 g/l glucose (PAA, Austria) supplemented with 10 % (v/v) heat-inactivated foetal bovine serum FBS (PAA, Austria)

and containing 100 U/ml penicillin / streptomycin (PAA, Austria). This basic culture medium will be referred to as complete DMEM in later sections. Cells were grown in 25, 75 or 175 $cm^2$ tissue culture flasks or in 10 or 15 cm tissue culture dishes in 37 ºC incubators and 5 % $CO_2$ atmosphere. In order to maintain a normal growth of these cells, they were kept sub-confluent at all times by passaging them routinely when they were 80-90% confluent. To do so monolayers were first washed with Dulbecco's PBS and then treated with 0.1 % trypsin (diluted in PBS/ EDTA) at 37 ºC for 5 min. Once cells were detached from the culture surface, they were subsequently collected in complete DMEM, spun (350 X *g*, for 5 min) and diluted 1:10 in fresh complete DMEM and transferred to new tissue culture flasks.

Murine BALB/c 3T3 wt and *scid* fibroblast cell lines were maintained in complete DMEM, but they were passaged every 3 days and 3 X $10^5$ cells were transferred to a new flask for each passage (Tordaro and Green, 1964). m*Tert* murine BALB/c 3T3 *scid* 3T3 fibroblasts were maintained likewise but in the presence of 3 µg/ml puromycin.

wt and *scid* hematopoietic stem cells were grown without passaging in serum free StemSpan® medium (Stemcell Technologies, Canada) containing 100 ng/ml murine stem cell factor (SCF), 100 ng/ml murine Fms-related tyrosine kinase 3 (Flt-3), 20 ng/ml human interleukin-6 (IL-6) and 100 U/ml penicillin / streptomycin. Cells were grown in 24-well non tissue culture treated plates at 37 ºC incubators and 5 % $CO_2$ atmosphere.

#### 2.2.4.3 Long-term cell storage

For long term storage of cell lines a 90 % confluent monolayer from 75 $cm^2$ tissue culture flask or 2-5 X $10^6$ cells were pelleted by centrifugation at 350 X *g* for 5 min, resuspended in 1 ml of freezing medium (50 % FCS, 40 % DMEM and 10 % Dimethyl sulfoxide (DMSO)) and transferred to a cryovial. Cells were frozen slowly overnight to -80 °C in an isopropanol freezing box and then transferred into liquid nitrogen.

To retriever frozen cells, aliquots were thawed rapidly in a 37 °C waterbath, transferred to 9 ml complete DMEM medium, spun at 350 X *g* for 5 min to remove the DMSO and finally resuspended in complete medium and cultured normally.

### 2.2.5 Plasmids transfection

In this thesis, the plasmids were transfected using the calcium phosphate coprecipitation method. In 10-cm tissue culture plates, 1 X $10^6$ BALB/c 3T3 *scid* m*Tert* fibroblasts were seeded and allowed 4 h to attach to the tissue culture surface.

10 µg of plasmid DNA (or a maximum of 30 µg, in cases where more than one plasmid was used) was placed in a 50 ml conical tube. Next, the DNA was mixed with 50 µl of 2.5 M $CaCl_2$, vortexed and then incubated for 5 min at RT. To each tube, 500 µl of 2 X HBS-buffered saline [281 mM NaCl, 100 mM HEPES, 1.5 mM $Na_2HPO4$ (pH 7.12)] was added drop-wise while vortexing at speed, and then immediately added to the plated cells and mixed. At about 14-16 h post-transfection, the medium was replaced after cells were twice washed with PBS. Cells and/or supernatants were harvested on days 2-3 post-transfection.

### 2.2.6 Western Blotting

#### 2.2.6.1 Total protein isolation

Total protein was isolated from cells transfected with plasmid DNA (section 2.2.5 above) or transduced with lentiviral vectors (Aiuti *et al*., 2007). Three days later, 5 X $10^5$ cells were washed twice in pre-cooled Dulbecco's phosphate buffer saline (PBS) (PAA, UK), then they were lysed in 200 µl of ice-cold protein lysis buffer, which consist of Radio Immuno-Precipitation Assay (RIPA) buffer (de Campos-Nebel *et al*., 2008) and freshly added protease inhibitor cocktail (Roche, UK). RIPA buffer was prepared using 150 mM NaCl, 1 % Triton X-100, 1 % sodium deoxycholate, 100 mM NaF, 0.1 % sodium dodecyl sulphate (SDS) and 50 mM Tris-HCl [pH 8.0]. Protease inhibitor cocktail tablets (containing serine, threonine, cysteine, aspartate and glutamic acid protease inhibitors) were prepared by dissolving 1 tablet in 1 ml of RIPA buffer to make 7 X protease inhibitor stock and then 1 X solution was made by diluting in RIPA buffer. Cells were harvested by detaching with a cell scraper, transferred into a 1.5 ml microfuge tube and centrifuged at 14000 X *g* at 4 °C to remove large debris, and finally samples were kept at -20 °C.

#### 2.2.6.2 Protein quantitation

Total protein extracted from cultured cells was quantified using Bio-Rad DC protein assay (Bio-Rad, USA) according to the manufacturer's instructions. The protein concentration was calculated from a standard curve, using a series of known Bovine serum albumin (BSA) standards concentrations.

#### 2.2.6.3 SDS-PAGE Gel casting

Discontinuous acrylamide gels of 12 % for lower (resolving) and 5 % for upper (stacking) gels, were prepared. The lower gel was prepared using 4 ml acrylamide (30 %) (Merck, Germany), 2.5 ml lower gel buffer [1.5 M Tris-HCl and 0.4 % w/v SDS, (pH 8.8)], 100 µl 10 % w/v APS, 5 µl TEMED, and made up to 10 ml with $ddH_2O$. The upper gel was prepared using 0.67 ml acrylamide (30 %), 2.5 ml upper gel buffer [1 M Tris-HCl and 0.8 % w/v SDS, (pH 6.8)], 100 µl 10 % w/v APS, 5 µl TEMED, and made up to 4 ml with $ddH_2O$.

A Minigel™ apparatus (Bio-Rad, USA) was assembled and the lower gel was cast first, leaving 1 cm distance at the top for the stacking gel. 15 min later the upper gel was poured into this 1cm space and $ddH_2O$ was added to the top of the gel to remove bubbles and level the top of the gel.

#### 2.2.6.4 Gel Running

Proteins were denatured by heating in a thermo-block at 95 °C for 10 min. Wells were flushed thoroughly with RIPA buffer before loading samples. Next, between 2–25 µg protein in 1 X RIPA buffer per sample was prepared and bromophenol blue was added to samples to 0.05 %. Pre-stained Novex® sharp protein standard (Life Technologies, UK) was loaded alongside the samples. Then, the tank was filled with 1 X migration buffer [25 mM Tris-HCl, 0.1 % w/v SDS and 192 mM glycine, (pH 8.3)] and the gel was electrophoresed at 200 V until the bromophenol blue run out of the gel.

#### 2.2.6.5 Electrotransfer

The glass plates were dismantled and the stacking upper gel cut off and discarded; then the separating gel was equilibrated for 10 min in cold (1 X) transfer buffer [4.8 mM Tris, 3.9 mM Glycine and 0.375 % SDS, (pH 8.3) with methanol freshly added at 20 %

(v/v)]. Proteins were blotted onto Hybond™ ECL™ nitrocellulose membrane (Amersham Biosciences/GE Healthcare, USA) using a BIO-RAD electrotransfer apparatus with (1 X) transfer buffer/methanol at 150 mA for 1.5 h. Later the membrane was washed with Tris buffered Saline plus Tween (TBS-T) [50 mM Tris-HCl, 150 mM NaCl and 0.05 % (v/v) Tween-20 (pH 7.5)]. Non-specific binding was blocked in TBS-T containing 5 % (w/v) non-fat milk powder for 2 h at RT or overnight at 4 °C with gentle agitation. Finally, the membrane was washed several times with TBS-T.

#### 2.2.6.6 Membrane hybridisation

I used two protocols to visualise protein bands blotted onto nitrocellulose membrane: enhanced chemiluminescence (ECL) and infrared imaging (Odyssey system).

##### 2.2.6.6.1 ECL development

All antibody solutions were made in blocking solution (5 % milk powder, 0. 1 % tween-20). The nitrocellulose blots were hybridised overnight at 4 °C with mouse anti-FLAG M2 monoclonal primary antibody (Stratagene, UK) at 1:1000 dilution. Subsequently, the membranes were washed 5 min X 3-4 times with TBS-T, and incubated with HRP-conjugated secondary antibody polyclonal goat anti-mouse immunoglobulin (Dako, UK) at 1:2000 dilution in blocking solution for 2 h at RT with gentle agitation. Then, the membranes were washed and treated with ECL reagent as directed (Thermo scientific, UK) for 1 min. Finally, membranes were exposed to film (Kodak Biomax light film, USA) and this was developed.

###### 2.2.6.6.1.1 Membrane stripping

For the purpose of re-probing with alternative antibody, the nitrocellulose membrane was stripped after ECL development. The membrane was washed first with TBS-T to remove any traces of ECL reagents. Then, the membrane was placed in stripping buffer [100 mM 2-mercaptoethanol, 2 % SDS and 62.5 mM Tris] at 80 °C with gentle agitation for 30 min. The blot was rinsed with dd$H_2O$ then TBS-T, twice each, to ensure removal of any traces of stripping buffer. Then the membrane was blocked in TBS-T 5 % milk powder for 1 hour. To normalise for gel loading blots were re-probed with a mouse monoclonal anti α-tubulin antibody (Sigma Aldrich) at dilution of 1:2000 for 2

hour at RT with agitation, incubated with HRP secondary antibody and then developed as described above.

#### 2.2.6.6.2 Odyssey development

This protocol used an Odyssey infrared imaging system (Li-cor Bioscience, Germany). The same anti-FLAG primary antibody was used (section 2.2.6.6.1 above), and the secondary antibody was IRDye® 800CW goat anti-mouse (Li-cor GmbH, Germany) at 1:2000 dilution for 2 h. The blots were then washed 5 min X 3-4 times with TBS-T. To normalise for gel loading, the blots were hybridised with rabbit polyclonal anti α-tubulin primary antibody (Abcam, UK) at 1:2000 dilution then Alexa Fluor® 680 Goat anti-rabbit secondary antibody (Invitrogen, USA) at 1:5000 dilution. Then the blots were scanned at 700 and 800 nm channels for α-tubulin and anti-FLAG, respectively.

## 2.2.7 Preparation of lentiviral vectors

### 2.2.7.1 Lentiviral vector production

Lentiviral plasmids were kindly donated by Prof. Luigi Naldini (Milano, Italy). Second- and third-generation human immunodeficiency virus-1 (HIV-1)-based vectors were produced by transient transfection of HEK-293T cells using the calcium phosphate coprecipitation method (Yanez-Munoz *et al*., 2006). 3 X $10^6$ cells were seeded in 15-cm plates in 25 ml complete DMEM, 72 h prior transfection. When the cell confluence reached ~80 % (~3 X $10^7$ cells), cells were ready for transfection and the medium was replaced with 20 ml fresh complete DMEM, 2 h prior transfection.

For preparation of third-generation vectors, HEK-293T were co-transfected with four plasmids in molar ratio of 1:1:1:2 (packaging: REV: envelope: transfer). To 50 ml tube, 12.5 µg of integration-proficient packaging plasmid (pMDLg/pRRE) or pMDLg/pRRE-intD64V (integration-deficient through a point mutation leading to D64V change, which inactivates the catalytic site of integrase), 7 µg of VSV-G pseudotyping plasmid (pMD2.VSV-G), 6.25 µg of REV plasmid (pRSV-rev) and 32 µg of transfer plasmid of choice, were added. The DNA was mixed with 125 µl of 2.5 M $CaCl_2$ [14.3 *g* of 97 % pure $CaCl_2$ in 50 ml $ddH_2O$], vortexed and then incubated for 5 min at RT. To each tube 1,250 µl of 2 X HBS-buffered saline [281 mM NaCl, 100 mM HEPES, 1.5 mM

$Na_2HPO4$ (pH 7.12)] was added drop-wise whilst vortexing. DNA mixture was then immediately mixed gently with the media on the plated cells

At 14-16 h post-transfection, the medium was replaced with 18 ml of complete DMEM. Subsequently, at 48 h post-transfection, media containing viral particles were harvested and media replaced with 18 ml complete DMEM. A further harvest was done at 72 h post transfection.

The harvested viral supernatant was cleared by centrifugation at 1000 X *g* for 10 min, at RT and then passed through 0.22 µm-pore-size filters (Nalgene, USA) to remove any suspended cells or debris. The filtered supernatant was then transferred to high-speed polyallomer centrifuge tubes (Beckman, USA) and concentrated by ultracentrifugation (50,000 X *g*, for 2 h, at 4 °C). Later, the supernatant was decanted; the pelleted viral particles were resuspended in 50 µl of serum free DMEM and transferred to 1.5 ml microfuge tubes. Viral suspensions were spun at 1000 X *g* for 10 min at 4 °C to remove debris, and then transferred to new microfuge tubes. Finally, the viral stock was adjusted to 10 mM $MgCl_2$, DNase I was added at 5 U/ml and the mixture was incubated at 37 °C for 30 min to remove any residual DNA. Finally virus suspensions were aliquoted and frozen at -80 °C.

Second-generation HIV-1-based vectors were produced by the method described above using three plasmids in molar ratio of 1:1:2 (packaging: envelope: transfer). 16.25 µg of integration-proficient packaging plasmid (pCMVΔR8.74) or integration-deficient (pCMVΔR8.74-int D64V), 7 µg of envelope pseudotyping plasmid (pMD2.VSV-G), and 32 µg of transfer plasmid of choice, were used for transfection of 293T cells.

#### 2.2.7.2 Lentiviral vector titration by flow cytometry

Lentiviral vectors bearing enhanced green fluorescent protein (eGFP) expression cassettes were titrated by transduction of HeLa cells with serial dilutions of vector stock, followed by flow cytometry analysis (Nightingale *et al*., 2006). 1 X $10^5$ HeLa cells were seeded per well in 6-well tissue culture treated plates. Next day, the medium was replaced with 1 ml of complete DMEM containing 16 µg/ml polybrene (Sigma Aldrich). For each lentiviral vector sample to be titrated, 1.2 µl of viral stock was ten-

fold serially diluted ($10^{-3}$, $10^{-4}$, $10^{-5}$ and $10^{-6}$) in 1.2 ml complete DMEM. 1 ml of each dilution was added to the cells and incubated for 3 days. Subsequently, the cells were harvested by trypsinization and then fixed in 1 % paraformaldehyde (PFA) in PBS.

The samples were analysed using a FACSCanto II flow cytometer and FACSDiva software (BD Biosciences, USA) and the viral titre was estimated using the percentage of eGFP positive cells. A non-transduced (mock) sample was used to determine the gating of negative population and highly positive samples were used to set manual compensation. The viable cell population was selected in the forward scatter (FSC) *versus* side scatter (SSC) plot and the positive population determined in the FL1 (eGFP) versus FL2 (red auto- fluorescence) plot. Those dilutions that gave a range of 1-10 % eGFP positive cells were used to determine the titre applying the following formula:

eGFP transducing units (TU)/ml=% green cells x$10^5$ (cells/well at the time of vector addition) X 1/vector dilution

The plasmid pRRLsc-CEW used to prepare lentiviral vectors coding for eGFP driven by CMV was donated by Prof. Luigi Naldini; while the plasmid pHR'sc-SEW used to prepare lentiviral vectors coding for eGFP driven by SFFV was kindly donated by Prof Adrian Thrasher (University College London).

#### 2.2.7.3 Lentiviral vector titration by Real time qPCR

All lentiviral vectors were titrated by real-time quantitative polymerase chain reaction (qPCR). For each sample two viral dilutions (1:2000 and 1:20000) were prepared in complete DMEM and HeLa cells were seeded in 6-well plates at a density of 1 X $10^5$/well. The following day, the medium was replaced with 1 ml of complete DMEM containing 16 µg/ml polybrene and 1 ml of viral dilution. For each titration one well of cells were mock-infected with all the same reagents but without the addition of virus. At 24 h post-transduction, the cells were trypsinised, spun (350 X *g* for 5 min), and resuspended in 200 µl PBS. Then, Genomic DNA was extracted from the cell suspension using Qiagen DNeasy® tissue kit (Qiagen, UK). Q-PCR was used to quantify human β actin gene copies and lentiviral late reverse transcript (LRT).

Two separate qPCR reaction were run per sample: SYBR green was used in one set of reactions as a fluorescent dye to measure cDNA copies of Bushman's LRT (Butler *et*

*al*., 2001) and, a Taqman-type fluorescent probe was used to measure β-actin housekeeping gene copy number (Marras, 2006). LRT DNA standards were made using the plasmid pRRLsin_CEW to give known copy numbers ($10^2$-$10^7$ copies/ reaction). DNA standards for the β-actin reaction were made using serial dilutions of HeLa cell genomic DNA. For this DNA from a known number of viable HeLa cells was extracted by DNeasy extraction kit (Qiagen) and then diluted to have $10^0$-$10^5$ cell equivalents/ reaction.

For the LRT reaction, in a 20 µl reaction volume, 10 µl of (2 X) ABI universal master mix (Bioline, UK), 0.06 µl of 100 µM of corresponding primers, 0.4 µl of SYBR green (Bioline, UK), 3.23 µl of $ddH_2O$ and 6.25 µl of the DNA samples (including those from mock infected cells), DNA standard, or water as a non-template control, were mixed together. DNA standards, water and mock were prepared in triplicate while DNA samples were amplified in duplicate.

The following primers were used (Sigma Aldrich):

LRT-forward 5'-TGTGTGCCCGTCTGTTGTGT-3'

LRT-reverse 5'-GAGTCCTGCGTCGAGAGAGC-3'

For the β-actin reaction, in a 20 µl reaction, 10 µl of (2 X) universal Taqman master mix (Roche, Germany), 0.06 µl of 100 µM of corresponding primers, 0.4 µl of 100 µM of β-actin probe (Applied Biosystems, USA), 4.84 µl of $ddH_2O$ and 5 µl of the DNA samples, DNA standard or water as a non-template control, were mixed together. Genomic HeLa DNA standards, water and mock were prepared in triplicate and DNA samples in duplicate.

The following primers were used (Sigma Aldrich):

β-actin-forward 5'-TCACCCACACTGTGCCCATCTACGA-3'

β-actin-reverse 5'-CAGCGGAACCGCTCATTGCCAATGG-3'

Actin probe 5'-(VIC)-ATGCCCTCCCCCATGCCATCCTGCGT-(TAMRA)-3'

The standard cycling for both reactions was started with initial heating at 50 °C for 2 min, then 95 °C for 10 min, 50 cycles of [95 °C 15 sec, 60 °C 1 min].

### 2.2.8 Immortalisation of BALB/c 3T3 *scid* fibroblasts

#### 2.2.8.1 Titration of puromycin resistance

In order to determine the minimum concentration of puromycin for selection of transduced fibroblast clones, BALB/c 3T3 *scid* fibroblasts were grown in increasing concentrations of puromycin with a range from 1- 10 µg/ml. Puromycin stock solution was prepared 1 mg/ml (w/v), filtered with 0.22 µm filter, and kept frozen at -20 $^{o}$C. In a 6-well plate, 5 X $10^4$ cells/well were seeded in complete DMEM. The next day, puromycin was added and cells were grown for another 3 days. When emerging puromycin-resistant cells started to be visible as a monolayer, incubation was terminated and colonies were stained with crystal violet.

#### 2.2.8.2 Crystal violet staining

Crystal violet (Sigma Aldrich) stock was prepared at 1 % (w/v) in ethanol, kept at RT and a working solution was made up by 10-fold dilution in dd$H_2O$. The colonies to be stained were first gently washed with PBS, and then fixed with 1 % w/v PFA PBS for 10 min. The colonies were washed again with PBS and then 1 ml of crystal violet working solution was added and left for 5 min. Next, crystal violet was discarded and wells were washed for a few seconds until excess crystal violet was removed. Finally, the plates were left open upside down to dry and then purple colonies were counted by eye making the final count as colonies/well.

#### 2.2.8.3 Preparation of retroviral vectors

The telomerase reverse transcriptase (*Tert)* gene encodes the catalytic subunit of mouse telomerase. Retroviral vectors encoding the mouse *Tert* (m*Tert)* gene expression cassette were produced by transient transfection of HEK-293T cells using calcium phosphate coprecipitation method (Swift *et al*., 2001). The pBABEpuro retroviral transfer plasmid encoding m*Tert* and *Pac* was kindly donated by Manuel Serrano (CNIO, Madrid, Spain). The other two plasmids required for retroviral vector production (Figure 2.1) were the packaging plasmid pGagpol (M57), and envelop plasmid pEnv (K73-Eco).They were a kind gift of Dr. Christopher Baum (Hannover Medical School, Germany). The HEK-293T cells were transfected using a similar

protocol to that for lentiviral vectors production (section 2.2.7.1 above), but with a plasmid molar ratio of 5:1:4 (packaging: envelope: transfer). At 48 and 72 h post-transfection, the filtered supernatant was directly treated with DNase and frozen at -80 $^{o}$ C without ultracentrifugation.

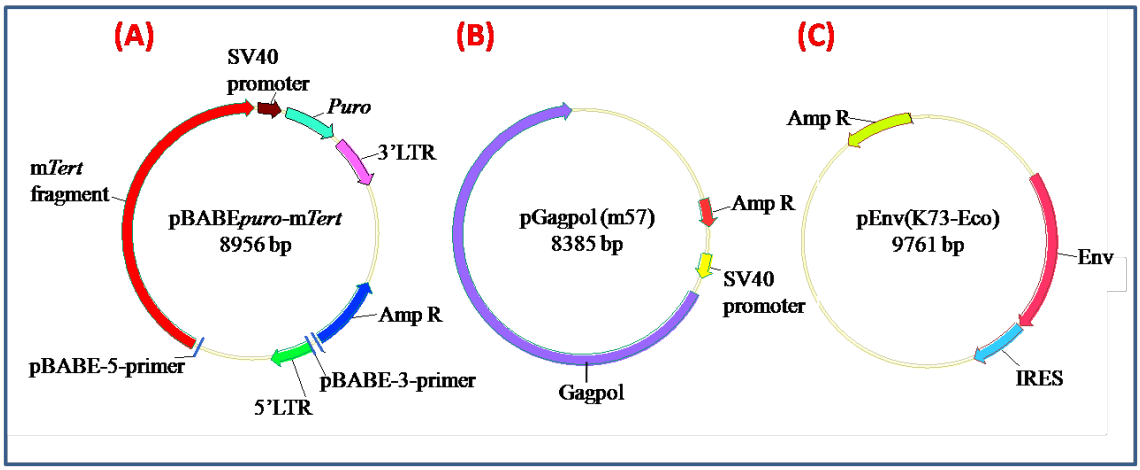



**Figure 2.1: Plasmid constructs used for retroviral production.** (A) Genetic map of retroviral construct coding for m*Tert*. (B)& (C) Genetic map of retroviral packaging plasmids. m*Tert:* mouse telomerase fragment, *Puro*: puromycin resistance gene, LTR: retroviral long-terminal repeat and Gag, Pol & Env: viral specific genes.

### 2.2.8.4 Titration of retroviral vectors

Retroviral supernatant was titrated by transduction of BALB/c 3T3 fibroblasts with serial supernatant dilutions, followed by scoring of puromycin-resistant colonies. A 6-well plate was seeded with 5 X $10^4$ BALB/c 3T3 *scid* fibroblasts per well. The following day, serial ten-fold dilutions ($10^0$ to $10^{-7}$) of freshly harvested retroviral supernatant were prepared in complete DMEM. The cell medium was replaced with 1 ml fresh DMEM and 1 ml corresponding dilution in presence of 8 µg/ml of polybrene. Two days post transduction, puromycin was added to a final concentration of 3 µg/ml. Three days later puromycin-resistant isolated colonies were visible. The colonies were then stained by crystal violet to facilitate counting of colony forming units (CFUs). To calculate the retroviral titre the following formula was used: retroviral vector titre (TU)/ml = number of CFUs X $10^5$ (cells/well at the time of vector addition) X 1/vector dilution.

#### 2.2.8.5 Transduction by retroviral vectors

To immortalise BALB/c 3T3 *scid* fibroblasts by transduction with m*Tert*-bearing retroviral vectors, a 75 $cm^2$ flask was seeded with 5 X $10^5$ cells in 10 ml complete DMEM. Next day, the culture medium was replaced with 10 ml of retroviral supernatant in the presence of 8 µg/ml polybrene. At 24 h post-transduction, the transduction was repeated again by replacing the retroviral supernatant with another 10 ml of retroviral supernatant from second viral harvest in the presence of 8 µg/ml polybrene. Next day transduced cells were selected by the addition of 3 µg/ml puromycin. The cells were maintained under puromycin selection, passaged as required, and some cells were kept frozen in liquid nitrogen. These immortalised fibroblasts will be referred to as BALB/c 3T3 *scid* m*Tert* fibroblasts in later sections.

#### 2.2.8.6 Clonal efficiency and growth curve

The clonal efficiency experiment was done as follows: in complete DMEM, a specific cell number of (10 or 100) per well of either BALB/c 3T3 *scid* fibroblasts or BALB/c 3T3 *scid* m*Tert* fibroblasts were seeded in 6 wells of 6-well plate. Over the next 4 days, when visible colonies started to emerge, the plates were washed with PBS and then stained with crystal violet and finally, CFUs were counted.

Similarly, to show that immortalised cells were growing more efficiently, the growth curve comparison between immortalised and non-immortalised cells was done as follows: in complete DMEM, 5 X $10^5$ per well of either BALB/c 3T3 *scid* fibroblasts or BALB/c 3T3 *scid* m*Tert* fibroblasts were seeded in 6-well plate. Then cells were washed with PBS trypsinised and then counted in triplicate at 24, 48, 72 and 96 h post-culturing.

### 2.2.9 Testing the efficiency of ZFN cutting

The mismatch-specific plant DNA endonuclease Surveyor® *Cel*-I mutation detection kit for standard gel electrophoresis (Transgenomic, USA) was used to detect known point mutations and polymorphisms in heteroduplex DNA. The key component of the kit is the Surveyor endonuclease which is a member of the CEL family of plant endonucleases (Oleykowski *et al*., 1998). This endonuclease is able to generate site-

specific DNA cleavage with high efficiency at sites of base substitution mismatch and other alterations (Yang *et al*., 2000). The Surveyor is able to cleave both strands of a DNA heteroduplex on the 3'-side of the mismatch site leading to recognition of DNA insertion/deletion (InDel) mismatches and all base-substitution mismatches (Qiu *et al*., 2004; Sokurenko *et al*., 2001). This assay will be referred to as Surveyor *Cel*-I assay in later sections.

The Surveyor assay protocol typically includes four main steps (for more details see Figure 4.14):

- Sample PCR amplification.
- Hybridization of mixed amplicons to produce heteroduplex and homoduplex DNA.
- Treatment with Surveyor endonuclease.
- Analysis of DNA fragments by PAGE.

#### 2.2.9.1 *Cel*-1sensitivity testing

For the purpose of finding out what was the minimum percentage of DNA mismatches that can be detected by this assay, control samples were prepared containing different combinations of BALB/c 3T3 wt and *scid* m*Tert* fibroblasts. The cells were cultured in complete DMEM, harvested, counted, and then mixed in the following combinations in total cell number of 1 X $10^6$ cells per combination (Table 2.1). Later genomic DNA was extracted by DNeasy® kit.

**Table 2.1: Designed control combination percentages**

| BALB/c type | Combination (%) | | | | | | |
|---|---|---|---|---|---|---|---|
| m*Tert scid* | 99 | 95 | 90 | 80 | 70 | 60 | 50 |
| Wt | 1 | 5 | 10 | 20 | 30 | 40 | 50 |

The following primers (Eurofins MWG, Germany) were selected using Primer3 (version 0.4.0) from *scid* mouse genome.

Forward primer: 5' -GCAGACAATGCTGAGAAAAGG- 3'

Reverse primer: 5' -GCACAAAACAGACAAGGGTGT- 3'

The PCR amplification should generate a 304 bp parental fragment (for DNA sequence please see Appendix, Box 1) and when cleaved by Surveyor's (at the *scid* point mutation mismatch) should give rise to two cleaved smaller products of 88 and 216 bp. The designed controls were tested using same protocol as in section 2.2.9.2 below.

### 2.2.9.2 ZFN-treated sample testing

DNA mismatches in designed controls or unknown samples treated with ZFNs were tested using the same protocol. 1 X $10^6$ cells were spun at 270 X *g* for 5 min, the supernatant was gently discarded, and genomic DNA was extracted by Qiagen DNeasy® kit. Then, 100 µl of TE buffer containing eluted genomic DNA was transferred into PCR-compatible tubes and incubated as follows: 68 ºC for 15 min; 95 ºC for 8 min; then held at 4 ºC. Next, 2 µl of DNA sample from the previous step, 5 µl of 10 X AccuPrime buffer II (Invitrogen, USA), 2.5 µl of each primer 10 µM, 0.2 µl (1 unit) of AccuPrime Taq DNA polymerase (Invitrogen, USA) and 40.3 µl sterile dd$H_2O$ were mixed for PCR. The cycling conditions were: 95 ºC for 5 min; 35 cycles of [95 ºC for 30 sec, 60 ºC for 30 sec, and 68 ºC for 40 sec]; 68 ºC for 2 min; hold at 4ºC. Subsequently, the amplification was verified by running 5 µl of PCR reaction on PAGE. For hybridisation, 6 µl of 1 X AccuPrime buffer II was mixed with 3 µl of PCR product and run in thermocycler as follows for annealing:

95 ºC for 10 min then 95 ºC to 85 ºC (-2.0 ºC/s)
85 ºC for 1 min then 85 ºC to 75 ºC (-0.3 ºC/s)
75 ºC for 1 min then 75 ºC to 65 ºC (-0.3 ºC/s)
65 ºC for 1 min then 65 ºC to 55 ºC (-0.3 ºC/s)
55 ºC for 1 min then 55 ºC to 45 ºC (-0.3 ºC/s)
45 ºC for 1 min then 45 ºC to 35 ºC (-0.3 ºC/s)
35 ºC for 1 min then 35 ºC to 25 ºC (-0.3 ºC/s)
25 ºC for 1 min and finally hold at 4 ºC.

The digestion reaction for Surveyor nuclease was set up as follows: 9 µl of reaction mixture (from step 2), 1 µl of Surveyor enhancer, and 1 µl Surveyor nuclease were mixed and incubated at 42 ºC for 60 min in the thermocycler.

Finally digested DNA samples were analysed by PAGE (section 2.2.2.2 above) using 3 µl of loading buffer mixed with products from (step 3) and then loaded and

electrophoresed. Images were captured under UV illumination using a Bio-Rad Gel Doc 2000 system and analysed by ImageJ/ V-2 software.

### 2.2.10 Gene targeting in BALB/c 3T3 *scid* m*Tert* fibroblasts

The plasmid containing *Prkdc-neo* template (9210 bp) p*Prkdc*Hind*Pml*I*neo*F (Figure 4.8 b) was designed to encode a *neo* cassette which should confer resistance to G418 in gene-targeted cells. Importantly, regardless whether site-specific gene targeting or random integration has occurred, positive-selected cells should generate visible colonies, which can be counted at the end of experiment.

Two different gene targeting experiments were done to target BALB/c 3T3 *scid* m*Tert* fibroblasts. Firstly, since there were two different ZFN plasmids available I set out to investigate which ZFN version (or combination of both ZFNs) was more efficient in terms of producing more G418 resistant colonies as an indicator of ZFN-induced DSBs. This was done by plasmid co-transfection of wild type *Prkdc-neo* corrective template of (9210 bp) plus plasmids encoding the two different ZFN monomers either alone or in combination.

To do this initial gene repair experiment: 1 X $10^6$ BALB/c 3T3 *scid* m*Tert* fibroblasts were seeded in 10 cm tissue culture plates using complete DMEM. The cells were allowed 4 hours for attachment, and then the cells were transfected with plasmids [for quantities please see Table 4.2] encoding for *Prkdc-neo* corrective template and/or ZFNs. Two days later, the cells were twice washed with PBS and incubated in complete DMEM containing 800 µg/ ml of G418 (Sigma Aldrich) for positive selection [G418 stock solution was 80 mg/ml (w/v) PBS]. The G418 minimum killing dose for these cells had been previously determined (Rafael Yáñez, personal communication). Five days post-transfection, DMEM medium was replaced with complete DMEM containing G418. After four days, visible colonies emerged which were stained with crystal violet and counted. Simultaneously before staining some of these colonies were picked and transferred into a new 96-well, then into a 24-well, and finally into 6-well plates to continue growing. Then they were harvested either to be stored frozen in liquid nitrogen or to be used for further gene targeting events analysis.

Secondly, having identified the best combinations of ZFNs (in previous experiment), it was important to verify the functionality of ZFN cassettes as they were expressing from

prepared ZFN lentiviral vectors. This was done by plasmid transfection of *Prkdc-neo* corrective template plus viral transduction of only efficient ZFN versions and combinations demonstrated in previous experiment. A 1 X $10^6$ BALB/c 3T3 *scid* m*Tert* fibroblasts were seeded in 10 cm tissue culture plates using complete DMEM. The cells were allowed 4 hours for attachment, and then transduced with 50 MOI of ZFN IPLVs. Later, at intervals of 12, 24, 48 or 72 h post-transduction, the cells were transfected with donor plasmid (containing *Prkdc-neo* corrective template of 9210 bp). Next, the same steps were followed as in previous experiment and crystal violet stained colonies were counted.

#### 2.2.10.1 Molecular analysis of G418r colonies

G418-resistant BALB/c 3T3 *scid* m*Tert* fibroblasts from previous experiments (gene-targeted by *Prkdc-neo* template and two ZFN plasmids) were further analysed by a PCR based assay to demonstrate gene targeting events. The genomic DNA was extracted by Qiagen DNeasy® kit. Then two primers (Sigma-Aldrich) were designed using Primer3 software (version 0.4.0). The forward primer was selected from the donor template plasmid (template-internal primer); whereas the reverse one was selected from the genomic DNA outside the homologous recombination region and very close to the expected end of the donor template (template-external primer), (see Figure 4.10 E for more details). The amplicon length was 1335 bp. The primers sequences were as follows:

Forward primer 5'- TCGCCTTCTTGACGAGTTCT-3'

Reverse primer 5'- TTTTCCCCCTCATGTCACTC -3'

For PCR amplification, 50 µl reactions were prepared containing 10 µl 5 X GoTaq® (Promega, UK) reaction buffer, 1.5 µl of 10 mM dNTPs, 4 µl of 25 mM $MgCl_2$, 1.5 µl of 10 µM of each forward and reverse primer, 200 ng DNA template (genomic DNA extracted from gene-targeted cells), and 0.5 µl of 5 u/µl GoTaq DNA polymerase. Thermo-cycling conditions were: initial denaturation at 95 °C for 2 min, then 35 cycles of 95 °C for 45 sec, 59 °C for 60 sec, 72 °C for 72 sec, followed by final extension at 72 °C for 5 min. Subsequently, PCR amplicons were purified by Qiagen purification columns from the reaction.

Vector NTI software (Invitrogen) was used to predict restriction enzymes that could cut within the PCR product to confirm the sequence of the amplicons. All those samples that showed amplification of the right size of (1335 bp) were digested with *Bgl*II restriction enzyme. Then, one sample that showed amplification of the right size was selected to be digested separately with *Hind*III, *Bgl*II and *Xho*I for further confirmation purposes.

### 2.2.11 Isolation of Murine Lin$^-$ HSCs

For the purpose of isolation of murine lin$^-$ (lineage negative) HSCs from both wild type and *scid* mice the following steps were conducted as previously described by Mikkola (Mikkola *et al*., 2000). Mice strains BALB/cOlaHsd as wild type or BALB/cJHan(tm)Hsd-Prkdcscid (Harlan, UK) as *scid* mutants were used. 6-week old mice were killed using a schedule 1 procedure. Femurs and tibias were collected, muscles were removed and the bone ends were cut off. 23G needles were used to collect the bone marrow by flushing the bone shafts using ice-cold serum-free DMEM. The cells were disaggregated by pipetting and then filtered using 30 µm nylon mesh filters (Millipore, USA) to remove cells clumps.

Subsequently, MACS® Linage cell depletion kit (Miltenyi Biotec, Germany) was used as directed by the manufacturer, to deplete lin$^+$ HSCs with a cocktail of biotinylated antibodies against a collection of "lineage" antigens (CD5, CD11b, CD45R, Anti-Gr-1 (Ly- 6G/C), 7-4, and Ter-119 antibodies) and Anti-Biotin microbeads. As a result, lin$^+$ cells were magnetically labelled and column-retained; whereas the lin$^-$ HSCs were eluted (Orlic, 2002).In brief, bone marrow cells were washed twice with cold buffer [2 mM EDTA 0.2 % (w/v) BSA PBS] spun at (300 X *g* for 10 min at 4 °C). Next, the cells were resuspended in buffer (40 µl per $10^7$ cells) and biotin antibody cocktail (10 µl per $10^7$ cells) and then incubated for 20 min at 4 °C. Later, buffer (30 µl per $10^7$ cells) and anti-biotin magnetic microbeads (20 µl per $10^7$ cells) were added and incubated for 30 min at 4 °C. The cells were washed, centrifuged, applied in the magnetic separation columns and finally collected.

Purified lin$^-$ HSCs were typically cultured in non- treated tissue culture 24-well plates at density of 7.5 X $10^4$ - 1 X $10^5$ cells/well using the culture conditions previously

mentioned in section 2.2.4.2 above. Lin$^-$ HSCs were transduced with different LVs or adenoviral vectors (at different MOIs) and harvested after 72 h to be tested with *Bsa*WI assay and/or deep sequencing, or analysed by flow cytometry to detect eGFP positive cells. Their viability was checked by staining with DAPI stain (1 µg/ml final concentration) followed by analysis by flow cytometry at the same time of eGFP detection.

### 2.2.12 *Bsa*WI Assay

This experiment was designed to detect homologous recombination events in cells that were treated with ZFN and homologous *Prkdc* corrective template lentiviral vectors. A silent diagnostic restriction site of the enzyme *Bsa*WI has been introduced via site-directed mutagenesis (previously performed in Dr Yáñez's laboratory) 2 nucleotides upstream from the *scid* point mutation in the *Prkdc* corrective template. The presence of *Bsa*WI sites in genomic DNA were assayed using PCR and Southern blotting protocols as described below steps.

#### 2.2.12.1 *Prkdc* template cloning

The plasmid pHind-TK*Bsa*WI encoding for a *Hind*III fragment (7573 bp) of mouse *Prkdc* gene, was used as a template to amplify by PCR the *Prkdc*-*Bsa*WI corrective template amplicon of 1626 bp. The following primers (Sigma-Aldrich) were selected by Primer3 software (version 0.4.0):

Forward primer 5'- **TGCCGGGACCC**AATGTTTAGTTTTTATGAGTGTTTTGC-3'

Reverse primer 5'- **AACGGGACCC**CAAGCCATCTCTCTAGCCCTAC-3'

The sequence **GGGACCC** was *San*DI recognition site while **TGCC** and **AAC** were added to facilitate the digestion. For PCR amplification, a 100 µl reaction mix containing 20 µl 5 X GoTaq® (Promega, UK) reaction buffer, 2 µl 10 mM dNTPs, 3 µl 25 mM $MgCl_2$, 6 µl of 10 µM of each forward and reverse primers, 200 ng DNA template, and 0.5 µl GoTaq DNA polymerase was prepared. Thermo-cycling conditions were: initial denaturation at 95 °C for 2 min, then 35 cycles of 95 °C for 60 sec, 74°C for 60 sec, 72 °C for 90 sec, and final extension at 72 °C for 5 min. Subsequently, PCR amplicons were retrieved from the gel and purified by Qiagen purification columns, and

then digested with *San*DI restriction enzymes (NEB, UK) to generate cohesive ends, at 37 °C for 15 h. The samples were purified again by Qiagen columns to remove the residual DNA small fragments. The amplified *Prkdc* template was then cloned in reverse orientation, to prevent any possibility of RNA splicing, into new lentiviral vector backbones encoding for *Prkdc* template alone or ZFN monomer and *Prkdc* template together. The sequence of that *Prkdc* template was verified by sequencing (for full sequence please see Appendix 1 Box 2). This corrective *Prkdc* template will be referred to as "*Prkdc-Bsa*WI template" in later sections.

#### 2.2.12.2 Transduction with lentiviral vectors

The lentiviral vectors were prepared and qPCR titrated as explained earlier (section 2.2.7 above) using ZFN/template or template alone backbones.10000 cells of BALB/c 3T3 *scid* m*Tert* fibroblasts or *scid* HSCs were transduced with several qPCR-determined MOIs and ratios of ZFNs: template. The cells were cultured for 10 days (fibroblasts) or 3 days (HSCs), then harvested and finally their DNA extracted by DNeasy® Qiagen kit.

#### 2.2.12.3 PCR of the targeted *Prkdc* template

Two primers (Sigma-Aldrich) were chosen, from mouse genome, outside the region of homology of the corrective template (external to the *Prkdc*- *Bsa*WI template) using Primer3 software (version 0.4.0).

Forward primer 5'-AACAATCCTCCTCCGAACCT -3'

Reverse primer 5'-TGGAGGTGGAAGAACCAAAC -3'

The extracted genomic DNA from targeted cells (section 2.2.12.2 ) was used to amplify a *Prkdc* amplicon of 1708 bp (for full sequence please see Appendix 1 Box 3). For PCR amplification, 5 µl 5 X Longamp Taq (NEB, UK) reaction buffer, 0.75 µl 10 mM dNTPs, 1 µl of 10 µM of each forward and reverse primers, 100 ng DNA template (genomic DNA), 1 µl (2.5 units) Longamp Taq DNA polymerase and up to 25 µl sterile nuclease-free water, were prepared. Thermo-cycling conditions were: initial denaturation at 95°C for 30 sec, then 35 cycles of 95 °C for 30 sec, 59 °C for 60 sec, 65 °C for 75 sec, and final extension at 65 °C for 10 min. Subsequently, PCR products were visualised by gel electrophoresis.

#### 2.2.12.4 Digestion with *Bsa*WI

The samples that were targeted with *Prkdc* template and then PCR amplified with template external primers, were digested with *Bsa*WI restriction enzyme at 60 °C for 16 h.

#### 2.2.12.5 Southern blotting

##### 2.2.12.5.1 Preparation of hot ladder

To set up hot (radioactively-labelled) ladder, 1 µg DNA ladder (HyperLadder I, Bioline), 2 µl 10x nick translation buffer (NTB) [0.5 M Tris HCl pH 7.5, 0.1 M $MgSO_4$, 1 mM Dithiothreitol (DTT)], 4 µl CGT [0.5 mM each dCTP, dGTP and dTTP in 10 mM Tris HCl pH 7.5], 1 µl (10 µCi) of [α -$^{32}$P] dATP (Perkin Elmer, USA), and 5 U Klenow (5 u/µl), were mixed together in a 20 µl reaction volume. The mixture was incubated at room temperature for 90 min, and then 40 µl of TE [10 mM tris-HCl, 1 mM EDTA pH 8.0] buffer was added and mixed. A 1 µl sample of labelled was counted in Biomax Geiger counter. Next, a MicroSpin S-300 column (GH Healthcare, UK) was packed and spun at 3000 X *g* for 1 min then samples were loaded into column and spun at 3000 X *g* for 2 min to remove any unincorporated radiolabel. Finally, 1 µl of eluate was counted in Biomax Geiger counter. 15-45 Geiger counts worth of ladder was loaded per gel lane. The remainder of hot ladder was kept frozen at -20°C.

##### 2.2.12.5.2 Probe Labelling

Two probes were used for southern blotting. Both probes were amplified from the same genomic *Prkdc* template (for sequence please see Appendix 1 Bbox 3), the first was full-length PCR amplified fragment of 1708 bp; the second was shorter and generated after digestion of the first one with *Bsa*WI and purification of the bigger fragment (1068bp). The longer probe (1708 bp) was prepared by PCR amplification with *Prkdc* template external primers (section 2.2.12.4 above). For the shorter probe: the 1708 bp *Prkdc* amplicon was prepared by PCR amplification with *Prkdc* template external primers and digested with *Bsa*WI, fragments were visualised by gel electrophoresis, and the 1068 bp fragment was cut from the gel and purified using Qiaquick® gel extraction kit.

100 ng of DNA probe to be labelled was transferred to 1.5 ml screw cap microfuge tube and made up to 4 µl with TE buffer. Then, 1 µl of 60ng/µl random hexanucleotides (dN6) (SIGMA Aldrich) was added. The samples were kept at 95 °C for 5 min in thermo-block and transferred onto ice for 2 min and then spun down briefly. Next, 3 µl of 10x random priming (RP) buffer [0.5 M Tris HCl pH 6.9, 0.1 M MgSO4, 1 mM DTT], 3 µl of CGT [0.6 mM of each dCTP, dGTP and dTTP in 10 mM Tris HCl pH 7.5] and 12 µl of water were mixed together. Then 12 µl of this mixture was added to each probe sample followed by 2 µl (20 µCi) of [α-$^{32}$P] dATP. 1 µl Klenow (5 u/µl) was added and the mixture was incubated at RT, for 1 hour.

40 µl of TE buffer was added before 1 µl of the mixture was counted in Biomax Geiger counter. Subsequently, MicroSpin S-300 column were packed and spun at 3000 X *g* for 1 min and then samples were loaded into column and spun at 3000 X *g* for 2 min. Finally, a 1 µl of eluate was assayed in a Biomax Geiger counter to check the radio nucleotide incorporation efficiency.

#### **2.2.12.5.3** Gel preparation

A 0.7 % agarose gel was cast in 1 X TAE with final EtBr concentration of 0.5 µg/ml, and the lanes were loaded with *Bsa*WI- digested DNA (section 2.2.2.1 above) or hot ladder. Then, the gel was run in TAE-EtBr over night in cold room at 35 V constant voltage. Next day, a photo was taken and image saved using a Bio-Rad Gel Doc 2000 system.

#### **2.2.12.5.4** Membrane Transfer and UV Fixation

The gel was agitated in 0.25 N HCl for 10 min, washed twice with water, agitated in denaturing solution [0.4 N NaOH and 0.6 M NaCl] for 30 min, then washed twice with water, and agitated in neutralising solution [1.5 M NaCl and 0.5 M Tris HCl pH 7.5] for 30 min. Next, GeneScreen Plus membrane (Perkin Elmer, USA) was cut, pre-wetted in water for a few seconds and was equilibrated in 10 X SSC [1.5 M NaCl, 0.15 M trisodium citrate dihydrate] for 15 min. Later, a capillary blot was set up using 10 X SSC buffer, and transfer was left overnight. Next day, transfer was dismantled. The membrane was agitated in 0.4 N NaOH for 1 min and then agitated in 0.2M Tris-HCl, 1 X SSC, pH 7.5 for 1 min. Subsequently, the membrane was rinsed with 2xSSC and then fixed in UV cross-linker (UVP, USA).

#### 2.2.12.5.5 Hybridisation

The membrane was kept wet throughout hybridisation, washing and exposure steps. For hybridisation, first the membrane was wetted in 2 X SSC, spread in hybridisation tube and the 2 X SSC was drained out from tube. Then, 15 ml of Church mix [1 % (w/v) BSA, 7 % (w/v) SDS in 0.5 M phosphate buffer] was added to hybridisation tube. The hybridisation oven (Techne, UK) was pre-warmed by rotating for at least 1 hour at 68 °C. Next, the labelled probe was boiled at 95 °C for 5 min, then transferred onto ice for 2 min and added to the membrane inside a hybridisation tube. The hybridisation was left overnight at 68 °C, and next day the membrane was washed several times with 0.5 % (w/v) SDS in 2 X SSC at 65 °C. Finally the membrane was sealed in a bag and exposed to phosphoimager screen and scanned by Typhone-8600 scanner (Amersham Pharmacia Biotech, UK) and analysed by ImageJ/ V-2 software.

### 2.2.13 DNA-PKcs activity assay

Gene repair of mutated *Prkdc* will lead to restoration of normal gene function of DNA-PKcs activity. To assay for restoration of DNA-PKcs in gene repaired cells, the SignaTECT® DNA-dependent protein kinase assay system (Promega, USA) was used to quantitate DNA-PK activity in purified nuclear cell extracts. The assay quantifies the phosphorylation of a biotinylated p53-derived peptide substrate captured onto a streptavidin coated membrane. What is important is that the streptavidin matrix provides rapid and quantitative capture of biotinylated substrate molecules, based on the strong affinity of biotin for streptavidin (Goueli, 1997; Loong *et al*., 2004).

#### 2.2.13.1 Extraction of nuclear protein

The cells that were corrected by gene repair via ZFNs and corrective *Prkdc* template subsequently had their DNA-PK activity assessed. The nuclear protein extraction protocol was based on CelLytic Nuclear Extraction kit (Sigma Aldrich, USA) but all the required buffers were prepared in the laboratory. To perform nuclear protein extraction the cells were allowed to swell with hypotonic buffer, their plasma membranes disrupted, the cytoplasmic fraction was removed, and the nuclear proteins were released from the nuclei by a high salt buffer.

Briefly, *scid* corrected cells were collected, washed with PBS, and lysed with hypotonic lysis buffer [10 mM HEPES, pH 7.9, with 1.5 mM $MgCl_2$ and 10 mM KCl] or isotonic lysis buffer [5 mM Tris HCl, pH 7.5, with 2 mM $MgCl_2$, 3 mM $CaCl_2$, and 300 mM Sucrose] for m*Tert* fibroblasts or HSCs, respectively as the HSCs were more delicate cells. In both lysis buffers were included 1 mM DTT and 10 % (v/v) protease inhibitor cocktail (Roche, UK). Next, the samples were incubated on ice for 30 min, later 10 % IGEPAL (Sigma Aldrich) was added and then vortexed and immediately spun at 11000 X *g* for 30 sec at 4 °C.

Subsequently, the samples were resuspended in Extraction buffer [20 mM HEPES, pH 7.9, with 1.5 mM $MgCl_2$, 0.42 M NaCl, 0.2 mM EDTA, and 25 % (v/v) Glycerol] with 1 mM DTT and 5 % (v/v) protease inhibitor cocktail; and then kept on ice for 30 min with agitation. Next, they were spun at 11000 X *g* for 6 min at 4 °C, transferred into clean tubes and snap-frozen in liquid nitrogen and then kept at -80 °C. The final protein content was quantified by Bradford protein assay (Bio-Rad, UK).

#### 2.2.13.2 DNA-PKcs assay protocol

The nuclear extract was used as a source for the DNA-PKcs assay and the samples were tested in the presence of activation buffer which includes calf-thymus DNA and corrected for background phosphorylation by nonspecific kinases in the absence of calf thymus DNA. Enzyme samples were serially diluted in a two-fold dilution series down to 1:16 (enzyme: diluent) using 1 X enzyme dilution buffer and 0.1 mg/ml BSA. The ATP mix was prepared using 5.0 µl ATP and 0.05µl [$\gamma$-$^{32}$P] ATP (Perkins Elmer, USA) per reaction.

Then, for each sample reaction with activator 2.5 µl activation buffer, 5.0 µl 5 X reaction buffer, 2.5 µl biotinylated peptide substrate, 0.2 µl BSA (10 mg/ml) and 5.0 µl ATP mix were added. Whereas for each sample reaction without activator: 2.5 µl control buffer, 5.0 µl 5 X reaction buffer, 2.5 µl biotinylated peptide substrate, 0.2 µl BSA (10 mg/ml) and 5.0 µl ATP mix were added. Next, the reaction was initiated by adding the appropriate amount of enzyme sample, adjusted to a final volume of 25 µl using deionised water and incubated at 30 °C for 5 min.

The reaction was then terminated by adding 12.5 µl of termination buffer. Subsequently, 10 µl of each terminated reaction was spotted onto the SAM 2 membrane which was

washed once for 30 sec and three times for 2 min, with 2 M NaCl; then four times for 2 min with 2 M NaCl in 1 % $H_3PO_4$ and finally twice for 30 sec with deionised water. The SAM2 membrane was left to dry and incorporation of [γ -$^{32}$P] into the biotinylated peptide was scanned by phosphoimager (Typhone-8600 scanner, Amersham Pharmacia Biotech, UK) and confirmed by scintillation counting. Finally, *Prkdc*-mediated phosphorylation was calculated by subtracting the activity of the enzyme in the absence of activator (control buffer) from that of the enzyme in the presence of activator (activation buffer). The following formula was applied:

Enzyme specific activity in pmol ATP/minute/µg of protein = [(cpm reaction with activator – cpm reaction without activator) × (37.5)] / [(10) × (time min) × (µg of protein in reaction) × (specific activity of **[**γ-$^{32}$P**]** ATP)]

Where: 37.5 are the reaction volume in µl, 10 is the volume in µl of the sample, time in min was (5), and (specific activity of **[**γ-$^{32}$P**]** ATP) was calculated as follows:

The specific activity of **[**γ-$^{32}$P**]** ATP in cpm/pmol = [(37.5 ÷ 5) X] / 2,500

Where: 37.5 are the reaction volume in µl, 5 is the volume in µl of sample, X is the average cpm of the 5µl sample and 2,500 is the number of picomoles of ATP in the reaction. cpm is counts per minute.

### 2.2.14 Enrichment of corrected cells

Melphalan belongs to the class of nitrogen mustards that generate DNA interstrand cross-links, inducing DSBs that will be repaired by NHEJ pathway. Importantly, DNA-PK is crucial to the cellular response to this DNA damaging agents (Muller *et al*., 2000). Those cells that showed homologous recombination events detected by *Bsa*WI assay (section 2.2.12 above) were subjected to enrichment using the drug melphalan to further demonstrate their phenotype correction. First, the drug cytotoxic dose for BALB/c 3T3 wt and *scid* fibroblasts was detected using the MTT assay. Then corrected cells were enriched by culturing in presence of this optimised dose of melphalan.

### 2.2.14.1 MTT assay

The MTT (3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide) assay was used to detect the cell viability at different melphalan concentrations. The assay principle depends on the conversion of the water soluble MTT into insoluble MTT-formazan crystals by mitochondrial dehydrogenases of living cells. The crystals are then solubilised, and the concentration determined by optical density at 570 nm (van Meerloo *et al*., 2011). The MTT (Sigma Aldrich) stock solution was prepared in the dark at 4 mg/ml in PBS then filtered and kept at 4°C till use. In a 24-well plate, 1 X $10^4$ wt or *scid* m*Tert* fibroblasts were seeded in complete DMEM. After 24 h, the culture medium was replaced with fresh complete DMEM containing different melphalan concentrations (0.01 -50 µM). After one hour of incubation at 37 °C, medium containing the drug was replaced by fresh complete DMEM.

The plates were then incubated for 5 days prior to determine cell viability by measurement of MTT conversion. 200 µl MTT of stock solution were added to each sample, incubated for 4 h and then the medium removed gently without disturbing the formazan crystals. The plates were left for 1 hour to dry then DMSO was added to dissolve the crystals, and incubated with agitation for 10 min. Later, 100 µl from each sample were transferred into 96-well plate to measure the optical density at 570 nm using GloMax®- Microplate Multimode Reader (Promega, USA) The optical density in samples from cells with zero melphalan concentration was considered as 100 % viability.

### 2.2.14.2 Potential enrichment by melphalan

To verify that the corrected cells can be enriched by melphalan they were cultured under melphalan selection for three weeks. The stock solution of 150 mM melphalan (Sigma Aldrich) was prepared by dissolving 14 mg melphalan 95 % purity in 290.47 µl absolute ethanol acidified with 3 drops of concentrated hydrochloric acid. Triplicate cell samples of 1 X $10^5$ BALB/c 3T3 m*Tert scid* fibroblasts transduced with IDLVs encoding for SFFV ZFNs and corrective *Prkdc-Bsa*WI template were seeded in 9 cm dishes with complete DMEM. The following day, the optimised cytotoxic dose of melphalan (from above section 2.2.14.1) was added for 1 hour and then replaced with fresh complete

DMEM. The cells were kept in culture for 3 weeks with weekly melphalan exposure until the emergence of visible colonies. Finally, melphalan resistant CFUs were counted.

After counting, melphalan-enriched cells from previous part were collected and subjected to the *Bsa*WI assay (section 2.2.12 above) and compared alongside to non-enriched corrected cells to further demonstrate phenotype restoration.

### 2.2.15 *Ex vivo* studies

All *ex vivo* procedures in this study were performed at the Institute of Child Health (ICH) - UCL/ London and approved by UK Home Office regulations under project licence number 70/7024. Two mice strains (BALB/cOlaHsd) as wt and (BALB/ cJHan (tm)Hsd-Prkdcscid) as *scid* from Harlan, UK were used

Animals were housed in a specific-pathogen-free animal facility at ICH and fed sterile normal diet.

#### 2.2.15.1 Murine transplantation

Donor animals were either wt or *scid* males; whereas recipients were always *scid* females. The $lin^-$ HSCs were isolated as described earlier (section 2.2.11 above) then transduced where applicable with ZFN and corrective template lentiviral vectors by adding concentrated lentiviral stock directly to the growth media and the cells were transduced overnight at the conditions described in (section 2.2.4.2 above).

The following day the transduced cells were harvested for intravenous injection into irradiated mice. Before injections; recipients were sub-lethally irradiated with a total of 2.5 Gy/mice in a one single dose using caesium 137 gamma irradiation in IBL 437C0 irradiator (RAD SOURCE, USA). 1 X $10^6$ transduced $lin^-$ cells, resuspended in 200 µl PBS, were intravenously injected into the tail veins of irradiated *scid* mice using a 27 G bevel needle attached to a tuberculin syringe. This procedure was carried out by Dr Mike Blundell.

### 2.2.15.2 Animal bleeding

In order to analyse whether *ex vivo* correction of the *scid* mutation in lin-HSCs could result in immune reconstitution, animals were bled at weeks 9, 16 and 24 post-transplantation and peripheral blood cells were assayed by multicolour flow cytometric analysis for expression of lineage markers. At week 9 post-transplantation, blood samples were collected (~50 µl/ animal) from mice tail veins using a 27 G bevel needle attached to a tuberculin syringe. Then, hematopoietic cells were lysed in erythrocytes lysis solution to eliminate erythrocytes before labelling with specific antibodies. The lysis was performed by adding 2 ml/sample of 1 X RBC lysis buffer (Biolegend, USA), kept in dark at RT for 10 min and then spun at 300 X *g* for 10 min at RT. Later, the supernatants were discarded, the blood sample volumes were made up to 100 µl using sterile PBS. Then, the samples were stained with different monoclonal antibodies raised against mouse different differentiation markers by adding to each sample: 1 µl of each of anti-human/mouse CD45R (B220) PE antibody to target B lymphocytes, anti-mouse CD3e PerCP-Cy5.5 antibody to target T lymphocytes and anti-mouse CD11b APC to target myeloid cells (macrophages, natural killer cells and granulocytes). One mock unstained sample was also prepared. The samples were then mixed and kept for 30 min at 4 °C. Next, the samples washed with 3 ml of 0.5 % FCS-PBS and then spun at 300 X *g* for 10 min. Finally, the supernatants were discarded and samples were analysed by flow cytometry on CyAn™ ADP **(**Beckman Coulter, USA) and analysed with FlowJo Software (TreeStar Inc).

Control beads samples were prepared along with the prepared blood samples by adding one drop of each of positive and negative Copmbeads™ anti-mouse Ig, K/negative control (FBS) compensation particles set (BD Biosciences, USA) to 100 µl PBS. Then,1 µl of rat IgG2a K isotype control PE, Armenian hamster IgG isotype control PerCP-Cy5.5, or rat IgG2b K isotype control APC; or all of them together were added to beads samples, washed and prepared as above. Likewise, a 1 µl of anti-human/mouse CD45R (B220) PE antibody, anti-mouse CD3e PerCP-Cy5.5 antibody or anti-mouse CD11b APC were added either separately or together to another bead samples and preparation continued as above.

At weeks 16 and 24 post transplantation, all animals were bled again and samples were stained with same antibodies used in week 9 applying same protocol. Two additional antibodies were used to target mature T cells: anti-Mouse CD4 PE to target helper T cells and anti-Mouse CD8a APC to target cytotoxic T cells. Detection of any cell population of these developed T cells will be an indicator of *scid* mouse phenotype correction. At the same time wt female mouse of strain BALB/c OlaHsd, was bled and used as a control alongside with staining against CD45R, CD3e and CD11b markers.

# Chapter Three

# Cloning of ZFNs and homologous templates, and lentiviral production

# 3 Cloning of ZFNs and homologous templates, and lentiviral production.

This chapter describes the work undertaken to produce integration proficient lentiviral vectors (IPLVs) and integration deficient lentiviral vectors (IDLVs) expressing either ZFNs, wt corrective templates, or both of them, ensuring that ZFN genes are expressible and detectable. In order to complete this, ZFNs and homologous templates were cloned into lentiviral plasmid backbones as a first step to produce lentiviral vectors. Then, ZFN expression was confirmed by Western blotting of cell extracts obtained by transfection of ZFN constructs. Later in this chapter, ZFN IPLVs and IDLVs, homologous template IDLVs or ZFN/ homologous template IDLVs were produced. Other lentiviral vectors expressing eGFP reporter gene driven by either Cytomegalovirus (CMV) or Spleen focus-forming virus (SFFV) promoters were prepared for control/optimisation purposes. All prepared lentiviral vectors were titrated by qPCR and the eGFP vectors were additionally titrated by flow cytometry. Finally, ZFN expression from prepared lentiviral viral vectors was confirmed by Western blotting.

All ZFNs were received from Sangamo Biosciences Inc. (California-USA) as pVAX plasmid constructs. Two wt mouse genomic DNA sequences, long of 9210 bp (including non homologous sequence of 1637 bp encoding *neo* cassette) and short of 1626 bp, were used as donor corrective templates. Most of these DNA inserts were cloned into either pRRL or pCCL lentiviral backbones and only one construct was made in an intermediary plasmid. pRRL and pCCL are lentiviral transfer plasmids where the 5' HIV LTR promoter/enhancer has been replaced by either chimeric Rous sarcoma virus (RSV)-HIV or CMV–HIV LTRs, respectively. In pRRL, the enhancer and promoter from the U3 region of RSV are joined to the R region of the HIV-1 LTR. While in pCCL, the enhancer and promoter of CMV were joined to the R region of HIV-1. Additionally simian virus 40 (SV40) polyadenylation site and enhancerless origin of replication sequences are cloned downstream of the HIV LTR, replacing most of the human sequence remaining from the HIV integration site (Dull *et al*., 1998; Zufferey *et al*., 1998).

## 3.1 Cloning of plasmid constructs

New constructs cloned in current study have encoded different features. The main features and their importance are listed in Table 3.1below:

**Table 3.1: Main features of newly constructed plasmids**

| | |
|---|---|
| CMV | The immediate early cytomegalovirus promoter is widely used to obtain stable, reliable and high level expression of transgenes in mammalian cells (Williams *et al*., 2005). |
| SFFV | The Spleen focus-forming virus LTR promoter is commonly used for sustained and long term expression in hematopoietic cells (Linemeyer *et al*., 1982). |
| cPPT/ cTS | cPPT/cTS: central polypurine tract/central termination sequence, during reverse transcription and synthesis of the second strand of DNA; only part of the initial RNA (cPPT) remains bound to the first DNA strand. The cPPT serves as a primer for the extension of the second strand, and allows reverse transcription to switch strand synthesis. (Follenzi *et al*., 2000). |
| Gag | Gag proteins comprise the encapsidation signal which is essential for the assembly of viral particles to provide viral basic physical infrastructure. |
| PBS | Primer-binding site, will act as a primer, binds to the cellular tRNA during the process of reverse transcription (Wakefield *et al*., 1996). |
| RRE | Rev responsive element is the binding site for the Rev protein, which facilitates export of the mRNA from the nucleus in response to the Rev protein expression (Fankhauser *et al*., 1991). |
| RU5 and 3'dLTR | RU5 is a modified viral 5'dLTR containing the repeated R sequence and unique 5' (U5) region. Whereas 3'dLTR contains the unique 3' (U3) region in addition to a second copy of the 5'R region (Naldini *et al*., 1996). |
| SV40 | Simian virus-40 promoter/enhancer allows plasmid replication in cells that express T antigen. |
| WPRE | WPRE: Woodchuck hepatitis virus post-transcriptional regulatory element, its presence enhances a homogeneous length of mRNA polyA leading to improve nuclear export of mRNA which ultimately result in increase both lentiviral vectors titer and expression (Schambach *et al*., 2006; Zanta-Boussif *et al*., 2009). |
| ZFN1 | Zinc finger nuclease obligatory heterodimer 1 encoding for triple FLAG-NLS-17373-FokKK domain. |
| ZFN2 | Zinc finger nuclease obligatory heterodimer 2 encoding for triple FLAG-NLS-N2A-17834-FokEL2 domain. |

### 3.1.1 Cloning of wt *Prkdc* long homologous template

#### 3.1.1.1 Construction of pCCLsc *Prkdc*Hind*Bsa*WIF

The pCCLsc *Prkdc*Hind*Bsa*WIF plasmid was designed to encode the mouse *Prkdc* wt long corrective template with a unique site of *Bsa*WI that was introduced by site directed mutagenesis previously in Dr. Yáñez's laboratory. The genetic maps of the recipient transfer plasmid (pCCLsc ISce-IT), mouse wt template donor plasmid (p*Prkdc*Hind*Bsa*WI), and the new construct (pCCLsc *Prkdc*HindBsaWIF) are shown in Figure 3.1. The cloning was done by digesting p*Prkdc*Hind*Bsa*WI with *Xho*I and the resulting linear DNA was digested then with *Xma*I. The fragment of 5' *Xma*I_*Prkdc*Hind*Bsa*WI_ *Xho*I 3' (7389bp) was retrieved from a preparative gel and cloned into the vector pCCLsc ISce-IT, which was digested before using *Age*I and *Sal*I. The generated plasmid 13388 bp was verified by characteristic restriction using *Xmn*I generating fragment sizes of 4600, 3235, 2959 and 2294 bp; and *Sca*I generating fragment sizes of 6760, 5596, and 1032 bp.

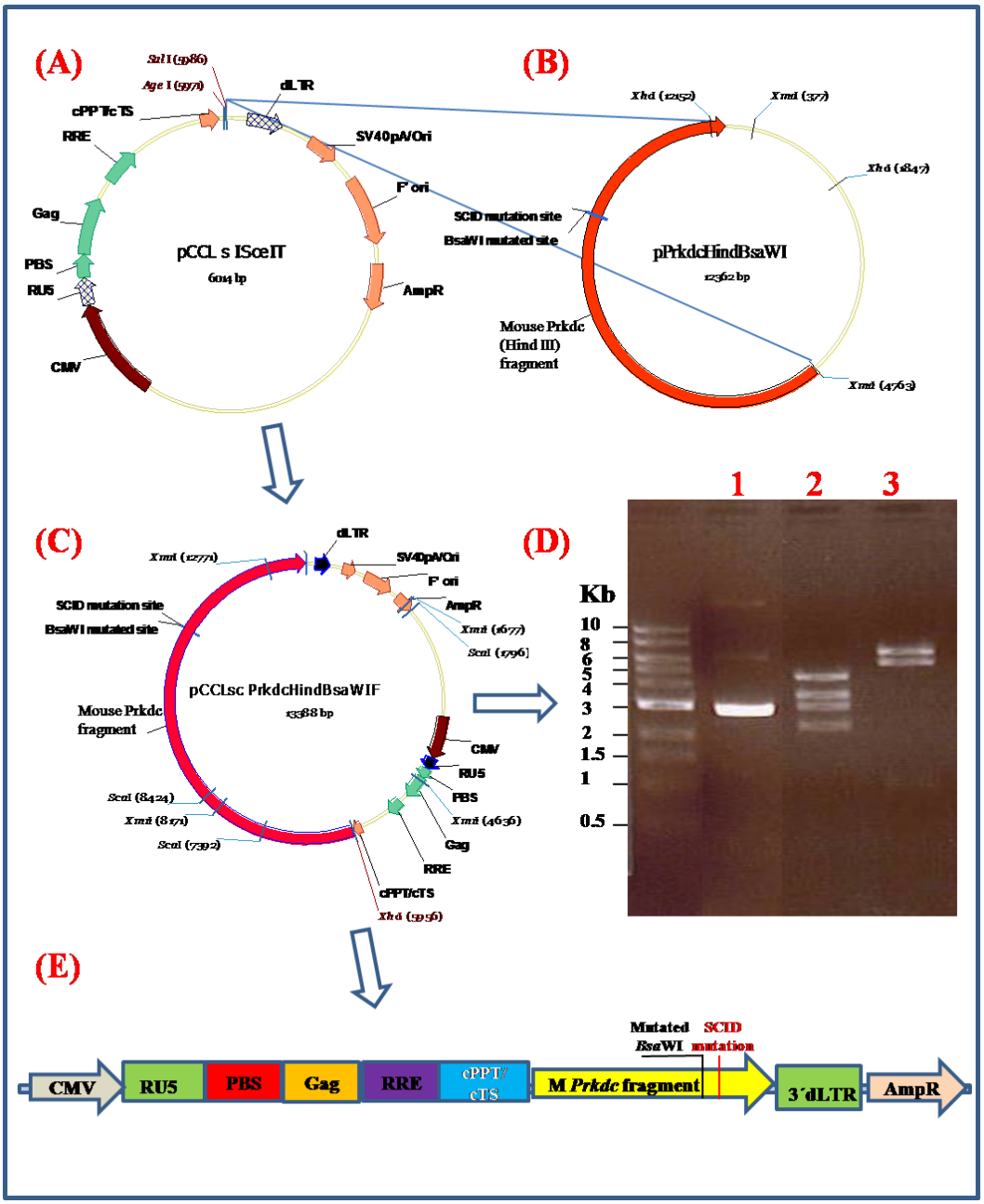



**Figure 3.1: Construction and characterisation of pCCLsc *Prkdc*Hind*Bsa*WIF.** (A), (B) and (C) show genetic maps of pCCLsc ISce-IT, p*Prkdc*Hind*Bsa*WI and pCCLsc *Prkdc*Hind*Bsa*WIF, respectively. (D) Restriction endonuclease analysis of generated pCCLsc *Prkdc*Hind*Bsa*WIF (13388 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Xmn*I generating fragments of 4600, 3235, 2959 and 2294 bp, and lane 3 digestion with *Sca*I generating fragments of 6760, 5596, and 1032 bp. DNA ladder: 5 µl of NEB 1 Kb. (E) A general schematic of pCCLsc *Prkdc*Hind*Bsa*WIF construct**.**

### 3.1.1.2 Construction of pCCLsc *Prkdc*Hind*Bsa*WIR

The genetic maps of recipient transfer plasmid (pCCLsc ISce-IT), mouse wt template donor plasmid (p*Prkdc*Hind*Bsa*WI), and the new construct (pCCLsc*Prkdc*Hind*Bsa*WIR) are shown in Figure 3.2. As in section 3.1.1.1 this plasmid was also designed to encode the mouse wt long homologous template containing a unique mutated site of *Bsa*WI, but in the reverse orientation. The cloning was done by double digestion of p*Prkdc*Hind*Bsa*WI with *Xho*I and *Xma*I. The fragment of 5' *Xma*I_*Prkdc*Hind*Bsa*WI_*Xho* I 3' (7389bp) was retrieved from a preparative gel and cloned into the vector pCCL s c ISce-IT, which was previously digested with *Age*I and then *Xho*I. The generated plasmid 13388 bp was verified by characteristic restriction using *Xmn*I generating fragment sizes of 4600, 3916, 2959 and 1913 bp and *Sca*I generating fragments sizes 9100, 3256 and 1032 bp.

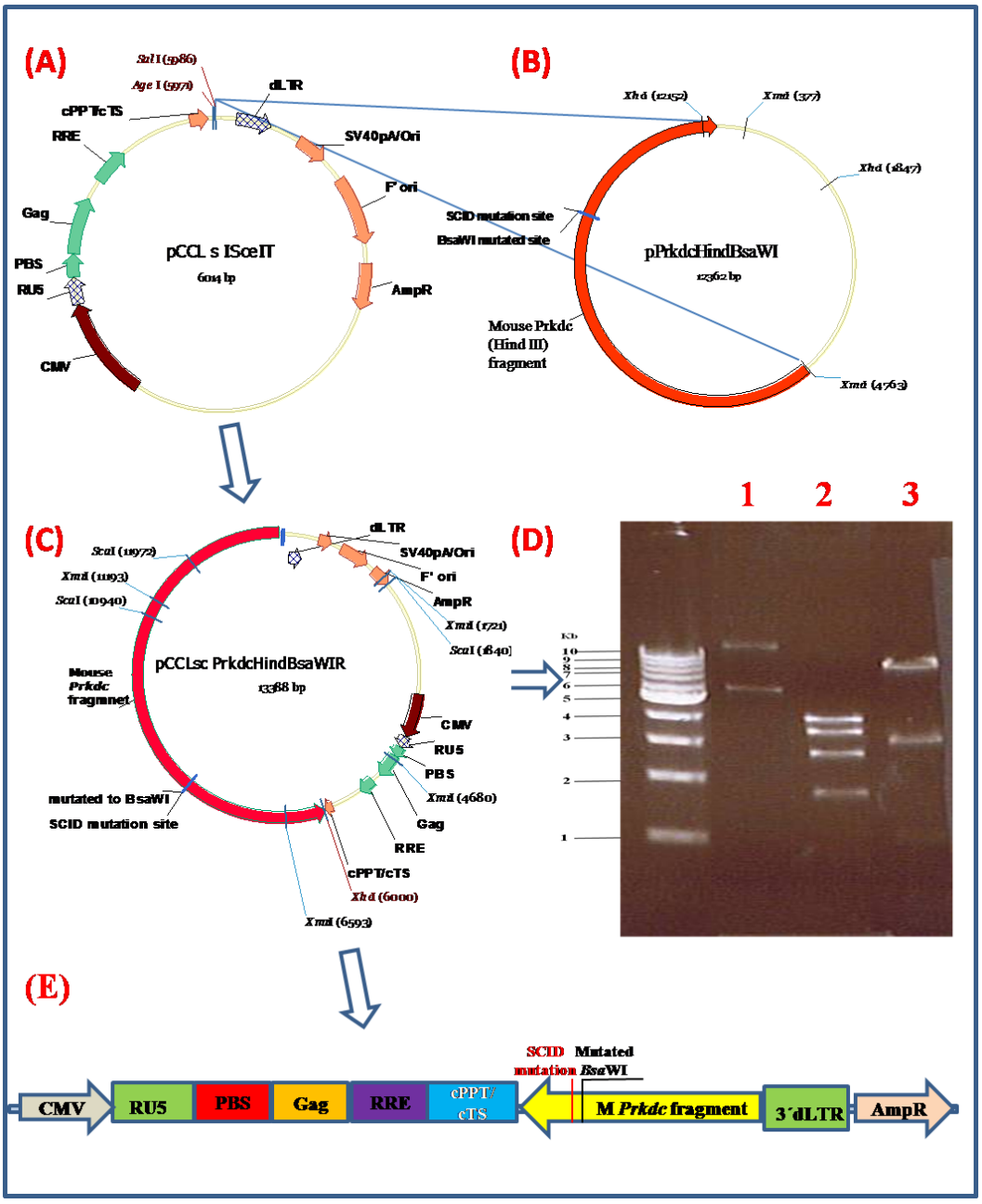



**Figure 3.2: Construction and characterisation of pCCLsc *Prkdc*Hind*Bsa*WIR.** (A), (B) and (C) show genetic maps of pCCLsc ISce-IT, p*Prkdc*Hind*Bsa*WI and pCCLsc *Prkdc*Hind*Bsa*WIR, respectively. (D) Restriction endonuclease analysis of generated pCCLsc *Prkdc*Hind*Bsa*WIR (13388 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Xmn*I generating fragments of 4600, 3916, 2959 and 1913 bp, and lane 3 digestion with *Sca*I generating fragments of 9100, 3256 and 1032 bp. DNA ladder: 5 µl of NEB 1 Kb. (E) A schematic of pCCLsc *Prkdc*Hind*Bsa*WIR construct.

### 3.1.2 Cloning of ZFNs

Three versions of *Prkdc* ZFNs were designed by Sangamo; the first was a first-screen product with monomers including wild-type *Fok*I domains which allow homodimerisation; the second and third versions included second-screen zinc-finger binding domains, linked to either wild-type or obligatory heterodimer *Fok*I domains. For all cloning strategies mentioned below, only the third version with obligatory heterodimeric ZFN monomers was used.

#### 3.1.2.1 Construction of pRRLsc_C_3FN-17373-FokKK_W

This plasmid was designed to encode the first ZFN obligatory heterodimer monomer (ZFN1). The genetic maps of the recipient transfer plasmid (pRRLsc_C_W), the ZFN1 donor plasmid (pVAX-3FN-17373-FokKK), and the new construct (pRRLsc_C_3FN-17373-FokKK_W) are shown in Figure 3.3. The construction of lentiviral transfer plasmid expressing ZFN1 under the control of CMV promoter was completed as follows: pVAX-3FN-17373-FokKK was digested with *Eco*RI, subjected to Klenow treatment for end-filling, and then digested using *Xho*I. The 5' *Eco*RI_3FN-17373-FokKK_*Xho*I 3' (1164 bp) fragment was retrieved from a preparative gel and cloned into the vector pRRLsc_C_W, which was previously digested with *Sal*I and *Eco*RV. The generated plasmid (7889 bp) was verified by characteristic restriction using *Xba*I generating a single fragment of 7889 bp, *Afl*II generating fragments of 4235 and 3654 bp and *Nco*I generating fragments of 5225, 1911 and 753 bp.

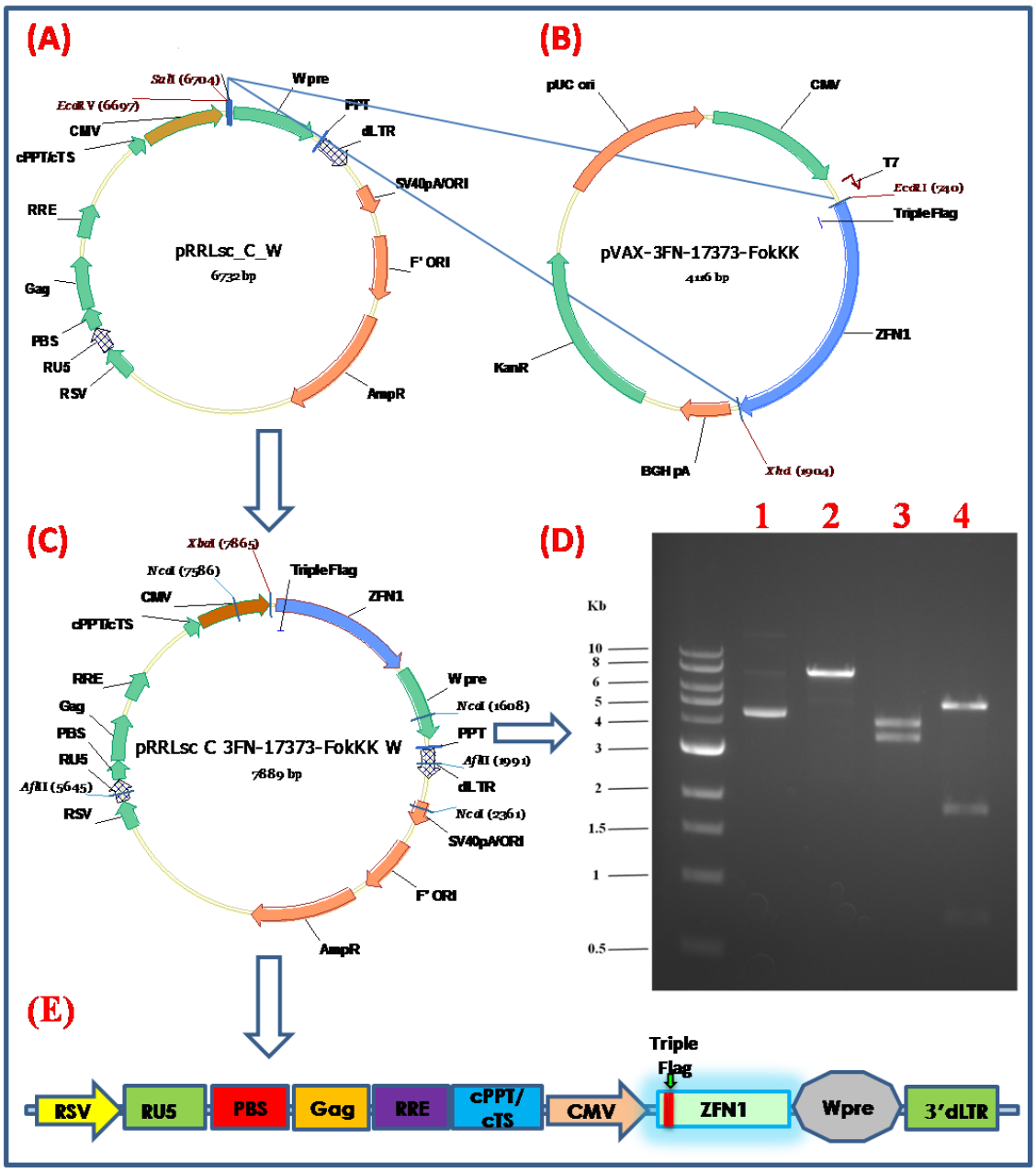



**Figure 3.3: Construction and characterisation of pRRLsc_C_3FN-17373-FokKK_W.** (A), (B) and (C) show genetic maps of pRRLsc_C_W, pVAX-3FN-17373-FokKK and pRRLsc_C_3FN-17373-FokKK_W, respectively. (d) Restriction endonuclease analysis of generated pRRLsc_C_3FN-17373-FokKK_W (7889 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Xba*I generating fragment size of 7889 bp, lane 3 digestion with *Afl*II generating fragment sizes of 4235 and 3654 bp and lane 4 digestion with *Nco*I generating fragment sizes of 5225, 1911 and 753 bp. DNA ladder: 5 µl of NEB I Kb. (E) Depicts a general schematic of pRRLsc_C_3FN-17373-FokKK_W construct.

#### 3.1.2.2 Construction of pRRLsc_C_N2A-3FN-17834-FokEL2_W

This plasmid was designed to encode the second ZFN heterodimeric monomer (ZFN2). The genetic maps of the recipient transfer plasmid (pRRLsc_C_W), ZFN2 donor plasmid (pVAX- N2A-3FN-17834-FokEL2), and the new construct (pRRLsc_C_N2A-3FN-17834-FokEL2_W) are shown in Figure 3.4. The construction of a lentiviral transfer plasmid expressing ZFN2 under the control of CMV promoter was completed as follows: pVAX- N2A-3FN-17834-FokEL2 was digested by *Eco*RI, then treated with Klenow fragment for end-filling, later was digested using *Xho*I. The 5' *Eco*RI_N2A-3FN-17834-FokEL2_*Xho*I 3' (1065bp) fragment was retrieved from preparative gel and cloned into the vector pRRLsc_C_W, which was previously digested with *Sal*I and *Eco*RV. The generated plasmid (7790 bp**)** was verified by characteristic restriction using *Xba*I generating fragment size of 7790 bp, *Kpn*I generating fragment sizes of 6202 and1588 bp and *Nco*I generating fragment sizes of 5225bp, 1501, 753 and 311 bp.

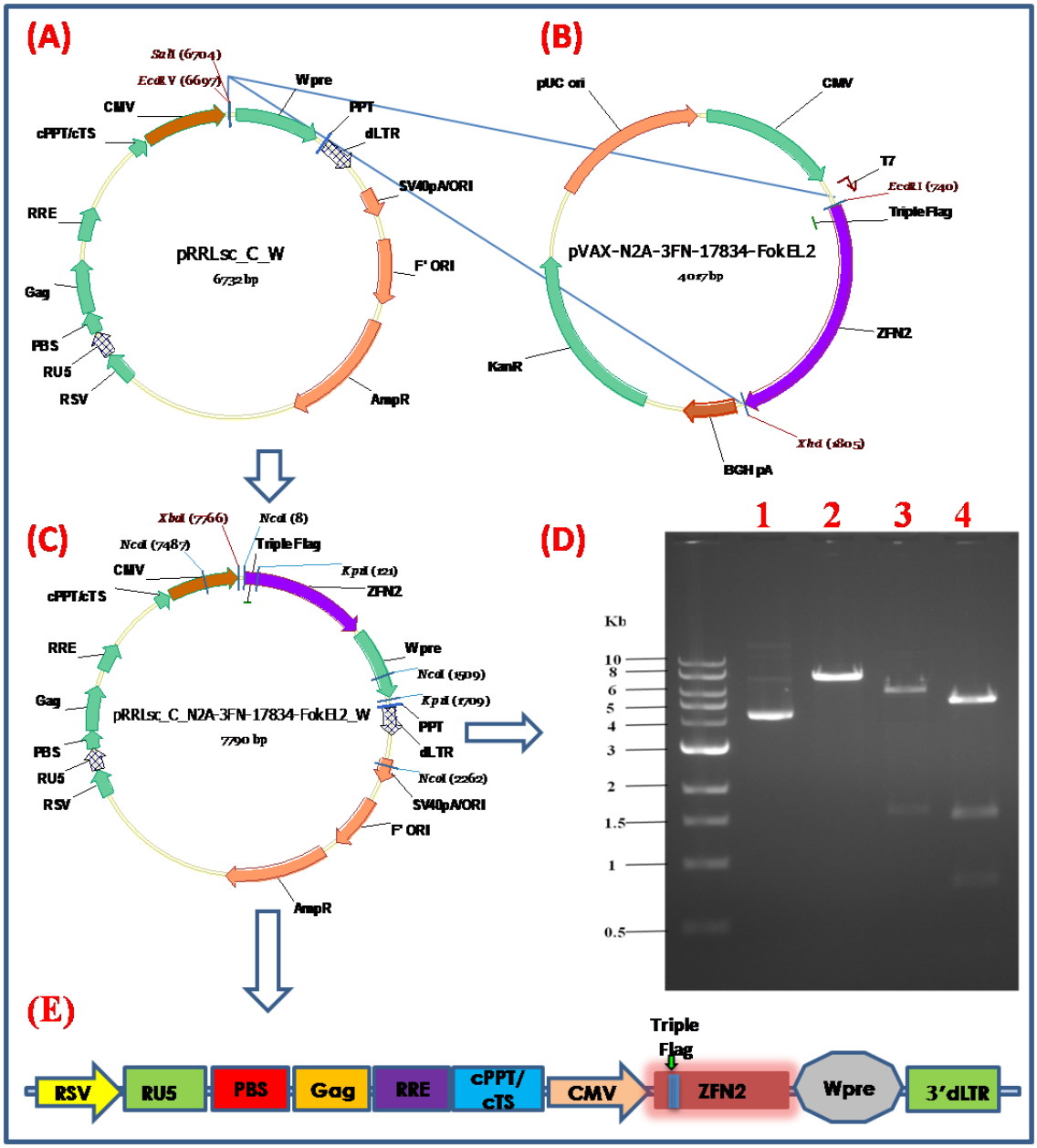



**Figure 3.4: Construction and characterisation of pRRLsc_C_N2A-3FN-17834-FokEL2_W.** (A), (B) and (C) show genetic maps of pRRLsc_C_W, pVAX- N2A-3FN-17834-FokEL2 and pRRLsc_C_N2A-3FN-17834-FokEL2_W, respectively. (D) Restriction endonuclease analysis of generated pRRLsc_C_N2A-3FN-17834-FokEL2_W (7790 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Xba*I generating fragment size of 7790 bp, lane 3 digestion with *Kpn*I generating fragment sizes of 6202 and1588 bp and lane 4 digestion with *Nco*I generating fragment sizes of 5225bp, 1501, 753 and 311 bp. DNA ladder: 5 µl of NEB I Kb. (E) A general schematic of pRRLsc_C_N2A-3FN-17834-FokEL2_W construct.

#### 3.1.2.3 Construction of pCCLsc_S_3FN-17373-FokKK_W

Here the aim was to construct a plasmid encoding one of the two ZFN heterodimerising monomers (ZFN1) under the control of SFFV promoter. The genetic maps of the recipient transfer plasmid (pCCLsc_S_W), ZFN1 donor plasmid (pVAX-3FN-17373-FokKK), and the new construct (pCCLsc_S_3FN-17373-FokKK_W) are shown in Figure 3.5. The construction of the lentiviral transfer plasmid expressing ZFN1 was completed as follows: pVAX-3FN-17373-FokKK was digested with *Eco*RI, ends were filled with Klenow, and then digested using *Xho*I. The 5' *Eco*RI_ 3FN-17373-FokKK_*Xho*I 3' 1164bp fragment was retrieved from a preparative gel and cloned into the vector pCCLsc_S_W, which was previously digested with *Age*I, treated with Klenow for end filling and then digested with *Sal*I. The generated plasmid (8248 bp**)** was verified by characteristic restriction using *Spe*I, generating a fragment of 8248 bp, *Eco*RI generating fragment sizes of 5934, 1791 and 523 bp, *Afl*II generating fragment sizes of 4097, 2191 and 1960 bp and *Nco*I generating fragment sizes of 4054, 3441 and 753 bp.

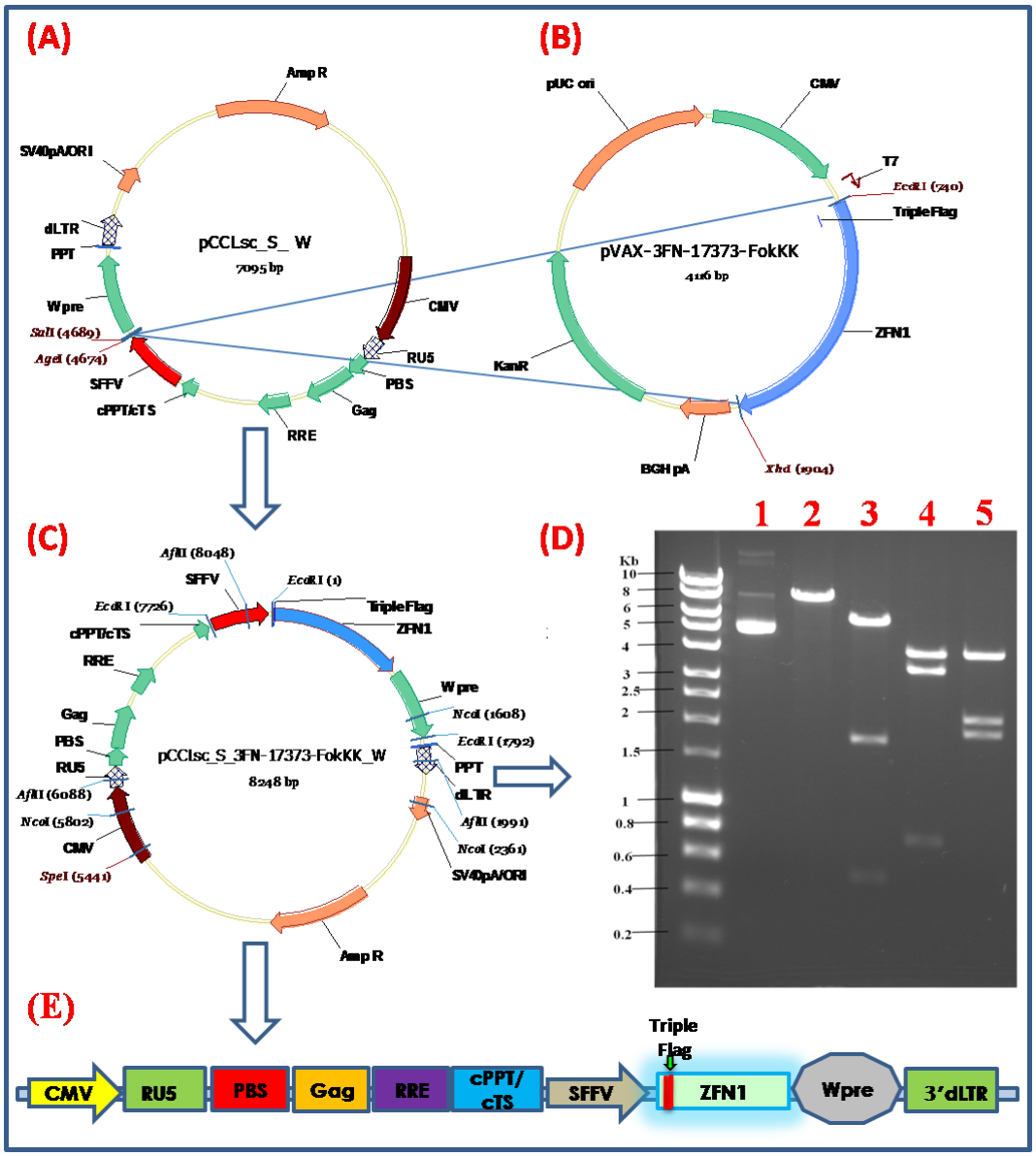



**Figure 3.5: Construction and characterisation of pCCLsc_S_3FN-17373-FokKK_W.** (A), (B) and (C) show genetic maps of pCCLsc_S_W, pVAX-3FN-17373-FokKK and pCCLsc_S_3FN-17373-FokKK_W, respectively. (D) Restriction endonuclease analysis of generated pCCLsc_S_3FN-17373-FokKK_W (8248 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Spe*I generating a fragment 8248bp, lane 3 digestion *Eco*RI generating fragment sizes of 5934, 1791 and 523 bp, lane 4 digestion with *Afl*II generating fragment sizes of 4097, 2191 and 1960 bp and lane 5 *Nco*I generating fragment sizes of 4054, 3441 and 753 bp. DNA ladder: 5 µl of Bioline Hyperladder I. (E) A general schematic of pCCLsc_S_3FN-17373-FokKK_W construct.

### 3.1.2.4 Construction of pCCLsc_S _N2A-3FN-17834-FokEL2_W

This plasmid was designed to encode the other ZFN heterodimeric monomer (ZFN2) under the control of the SFFV promoter. The genetic maps of the recipient transfer plasmid (pCCLsc_S_W), ZFN2 donor plasmid (pVAX- N2A-3FN-17834-FokEL2), and the new construct (pCCLsc_S_N2A-3FN-17834-FokEL2_W) are all shown in Figure 3.6. The construction of lentiviral transfer plasmid expressing ZFN2 driven by SFFV promoter was done as follows: pVAX- N2A-3FN-17834-FokEL2 was digested by *Eco*RI, the ends were filled using Klenow, then digested using *Xho*I. The 5' *Eco*RI_N2A-3FN-17834-FokEL2_*Xho*I 3' (1065bp) fragment was retrieved from preparative gel and cloned into the vector pCCLsc_S_W, which was previously digested with *Age*I, treated Klenow for end filling and then digested with *Sal*I. The generated plasmid (8149 bp**)** was verified by characteristic restriction using *Spe*I generating fragment size of 8149 bp, *Nhe*I generating fragment sizes of 6625 and 1524 bp, *Afl*II generating fragment sizes of 4097, 20924 and 1960 bp and *Nco*I generating fragment sizes of 3441, 2454, 1501 and 753 bp.

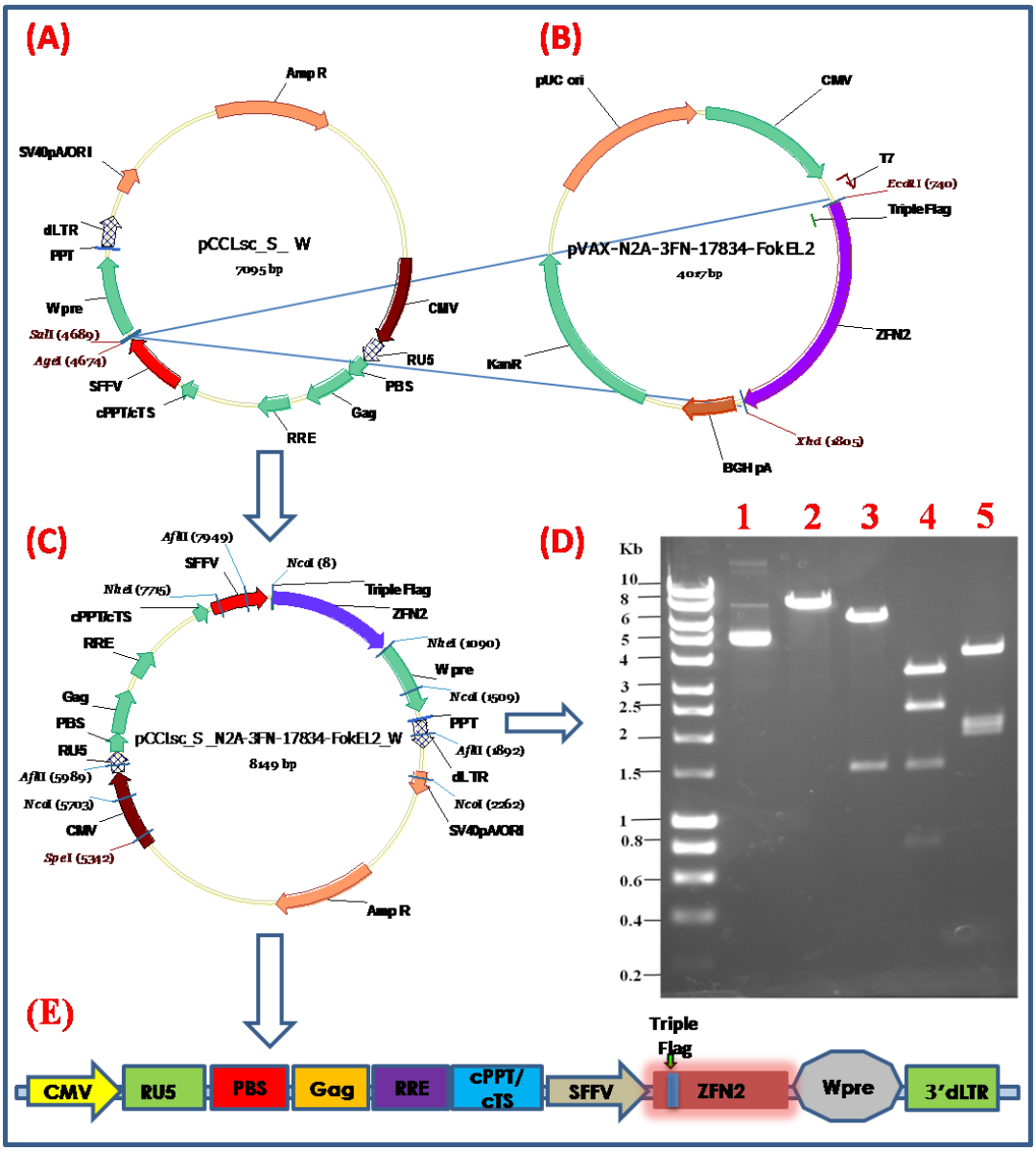



**Figure 3.6: Construction and characterisation of pCCLsc_S_N2A-3FN-17834-FokEL2_W.** (A), (B) and (C) show genetic maps of pCCLsc_S_W, pVAX-N2A-3FN-17834-FokEL2 and pCCLsc_S_N2A-3FN-17834-FokEL2_W, respectively. (D) Restriction endonuclease analysis of generated pCCLsc_S_N2A-3FN-17834-FokEL2_W (8149 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Spe*I generating fragment size of 8149 bp, lane 3 digestion *Nhe*I generating fragment sizes of 6625 and 1524 bp, lane 4 digestion with *Nco*I generating fragment sizes of 3441, 2454, 1501 and 753 bp and lane 5 *Afl*II generating fragment sizes of 4097, 20924, and 1960 bp. DNA ladder: 5 µl of Bioline Hyperladder I. (E) A schematic of the pCCLsc_S_N2A-3FN-17834-FokEL2_W construct.

### 3.1.2.5 Construction of pVAX-3FN-17834-FokEL2-3FN-17373-FokKK

For the main gene repair experiment, it is preferred to reduce the number of lentiviral vectors that co-transduce the targeted cells. In this regard, the expression of both ZFN monomers from one construct (with incorporation of a N2A self-processing polyprotein signal that allows monomers separation) would be better than express them from two different constructs. This construct was designed to be an intermediate plasmid encoding both ZFNs heterodimers (ZFN1+ZFN2) in a pVAX backbone which was then cloned into the lentiviral backbone. The genetic maps of recipient (ZFN1) transfer plasmid (pVAX-3FN-17373-FokKK), ZFN2 donor plasmid (pVAX- N2A-3FN-17834-FokEL2) and the new construct (pVAX-3FN-17834-FokEL2-3FN-17373-FokKK) are all shown in Figure 3.7. The construction of intermediate plasmid expressing both ZFN1 and ZFN2 was done as follows: pVAX-3FN-17373-FokKK was digested with *Bgl*II and *Xho*I; then the fragment of 5'-*Bgl*II-3Flag-NLs-17373-FokKK-*Xho*I-3' (ZFN1) of 1152 bp was retrieved from preparative gel and cloned into the plasmid pVAX- N2A-3FN-17834-FokEL2 (ZFN2) that was previously digested with *Bgl*II and *Xho*I. The generated plasmid (5157 bp**)** was verified by characteristic restriction using *Xho*I generating fragment size of 5157 bp, *BamH*I generating fragment sizes of 4011 and 1146 bp, and *Nco*I generating fragment sizes of 3198, 16246 and 335bp bp.

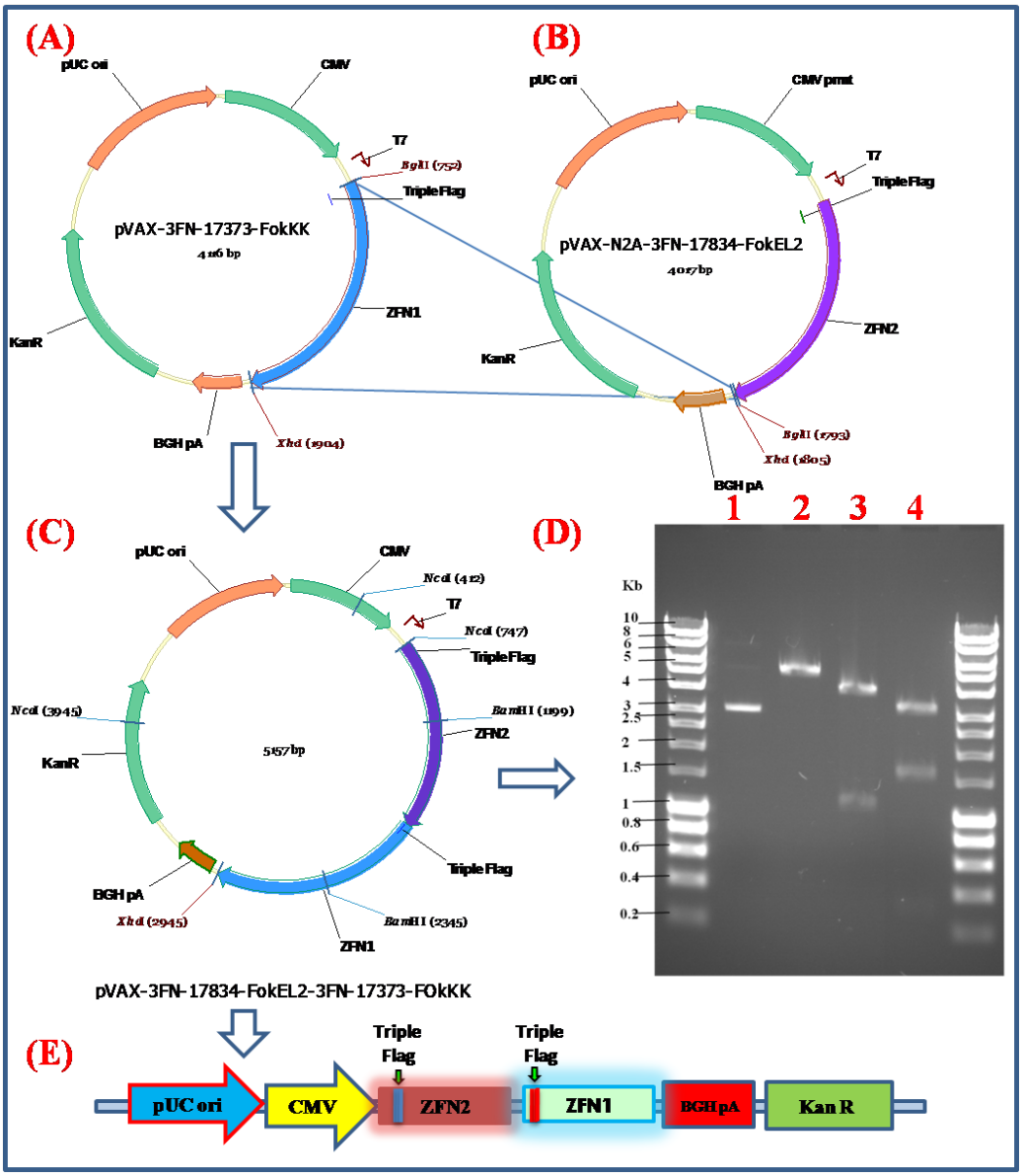



**Figure 3.7: Construction and characterisation of pVAX-3FN-17834-FokEL2-3FN-17373-FokKK.** (A), (B) and (C) show genetic maps of pVAX-3FN-17373-FokKK, pVAX- N2A-3FN-17834-FokEL2 and pVAX-3FN-17834-FokEL2-3FN-17373-FokKK, respectively. (D) Restriction endonuclease analysis of generated pVAX-3FN-17834-FokEL2-3FN-17373-FokKK (5157 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Xho*I generating fragment size of 5157 bp, lane 3 digestion *BamH*I generating fragment sizes of 4011 and 1146 bp and lane 4 digestion with *Nco*I generating fragment sizes of 3198, 16246 and 335bp bp. DNA ladder: 5 µl of Bioline Hyperladder I. (E) Depicts a schematic diagram of pVAX-3FN-17834-FokEL2-3FN-17373-FokKK construct.

#### 3.1.2.6 Construction of pRRLsc_C_3FN-17834-FokEL2-3FN-17373-FokKK_W

This plasmid was constructed to allow expression of ZFN1 and ZFN2 heterodimers simultaneously from the same RRL backbone under the control of the CMV promoter. Figure 3.8 below shows the genetic maps of the recipient transfer plasmid (pRRLsc_C_W), ZFN1 and ZFN2 donor plasmids (pVAX-3FN-17834-FokEL2-3FN-17373-FokKK) and the new construct (pRRLsc_C_3FN-17834-FokEL2-3FN-17373-FokKK_W). The construction of lentiviral transfer plasmid expressing both ZFN1 and ZFN2 was done as follows: pVAX-3FN-17834-FokEL2-3FN-17373-FokKK was digested by *Eco*RI, subjected to Klenow treatment for end-filling, and then digested using *Xho*I. The 5' *Eco*RI_3FN-17834-FokEL2-3FN-17373-FokKK_*Xho*I 3' (2205 bp) fragment was retrieved from a preparative gel and cloned into the plasmid pRRLsc_C_W previously digested with *Sal*I and *Eco*RV. The generated plasmid (8930 bp) was verified by characteristic restriction using *Age*I generating fragment size of 8930 bp, *Afl*II generating fragment sizes of 5276 and 3654 bp, *Bam*HI generating fragment sizes of 7306, 1146 and 478 bp and *Nco*I generating fragment sizes of 5225, 2641,753 and 311 bp.

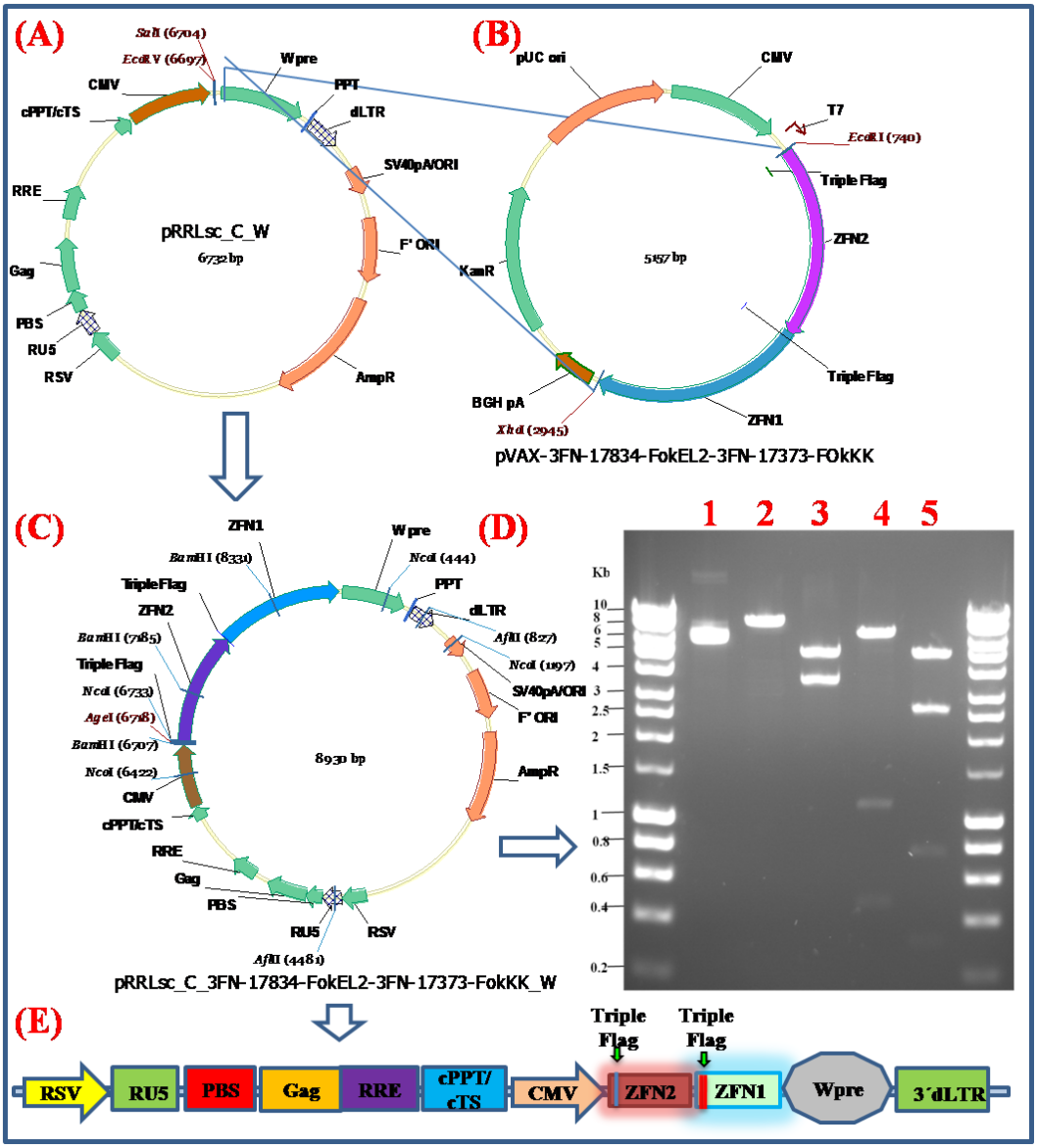



**Figure3.8: Construction and characterisation of pRRLsc_C_3FN-17834-FokEL2-3FN-17373-FokKK_W.** (A), (B) and (C) show genetic maps of (pRRLsc_C_W), (pVAX-3FN-17834-FokEL2-3FN-17373-FokKK) and (pRRLsc_C_3FN-17834-FokEL2-3FN-17373-FokKK_W); respectively. (D) Restriction endonuclease analysis of generated pRRLsc_C_3FN-17834-FokEL2-3FN-17373-FokKK_W (8930 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Age*I generating fragment size of 8930 bp, lane 3 digestion with *Afl*II generating fragment sizes of 5276 and 3654 bp, lane 4 digestion with *Bam*HI generating fragment sizes of 7306, 1146 and 478 bp and lane 5 digestion with *Nco*I generating fragment sizes of 5225, 2641,753 and 311 bp. DNA ladder: 5 µl of Bioline Hyperladder I. (E) Depicts a schematic diagram of the pRRLsc_C_3FN-17834-FokEL2-3FN-17373-FokKK_W construct.

#### 3.1.2.7 Construction of pCCLsc_S_3FN-17834-FokEL2-3FN-17373-FokKK_W

This plasmid was constructed to express ZFN1 and ZFN2 heterodimers simultaneously from same CCL backbone under the control of SFFV promoter. Figure 3.9 below shows the genetic maps of the recipient transfer plasmid (pCCLsc_S_W), ZFN1 and ZFN2 donor plasmids (pVAX-3FN-17834-FokEL2-3FN-17373-FokKK) and the new construct (pCCLsc_S_3FN-17834-FokEL2-3FN-17373-FokKK_W). The construction of lentiviral transfer plasmid expressing both ZFN1 and ZFN2 was done as follows: pVAX-3FN-17834-FokEL2-3FN-17373-FokKK was digested with *Eco*RI, subjected to Klenow treatment for end-filling, and then digested using *Xho*I. The 5' *Eco*RI_3FN-17834-FokEL2-3FN-17373-FokKK_*Xho*I 3' (2205 bp) fragment was retrieved from a preparative gel and cloned into the plasmid pCCLsc_S_W previously digested with *Age*I then treated with Klenow and finally digested with *Sal*I. The generated plasmid (9289 bp) was verified by characteristic restriction using *Spe*I generating fragment size of 9289 bp, *Pvu*I generating fragment sizes of 8238 and 1051 bp, *Bam*HI generating fragment sizes of 7669, 1146 and 474 bp and *Nco*I generating fragment sizes of 3441, 2641, 2454 and 753 bp.

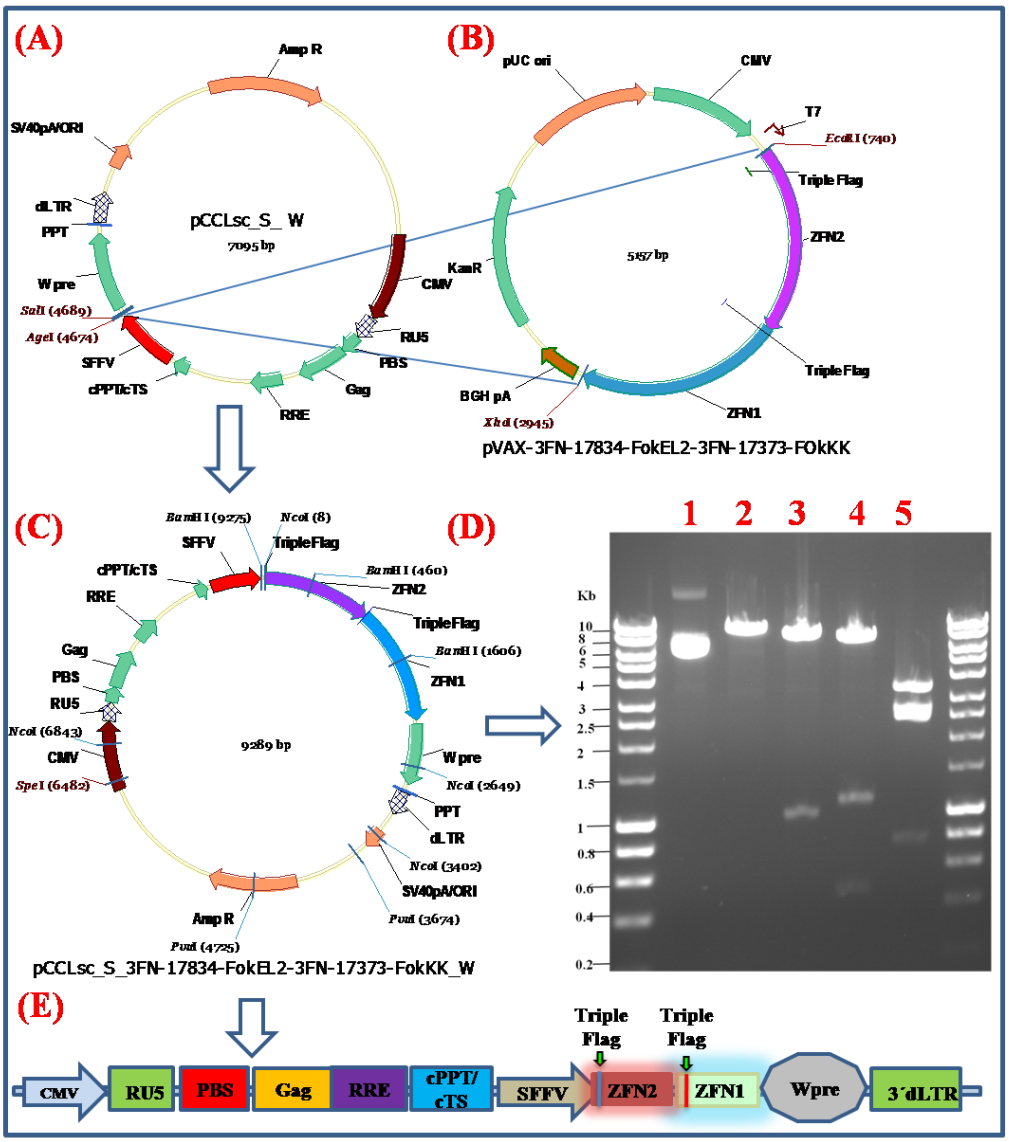



**Figure 3.9: Construction and characterisation of pCCLsc_S_3FN-17834-FokEL2-3FN-17373-FokKK_W.** (A), (B) and (C) show genetic maps of (pCCLsc_S_W), (pVAX-3FN-17834-FokEL2-3FN-17373-FokKK) and (pCCLsc_S_3FN-17834-FokEL2-3FN-17373-FokKK_W); respectively. (D) Restriction endonuclease analysis of generated pCCLsc_S_3FN-17834-FokEL2-3FN-17373-FokKK_W (9289 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Spe*I generating fragment size of 9289 bp, lane 3 digestion with *Pvu*I generating fragment sizes of 8238 and 1051 bp, lane 4 digestion with *Bam*HI generating fragment sizes of 7669, 1146 and 474 bp and lane 5 digestion with *Nco*I generating fragment sizes of 3441, 2641, 2454 and 753 bp. DNA ladder: 5 µl of Bioline Hyperladder I. (E) Depicts a general schematic diagram of the pCCLsc_S_3FN-17834-FokEL2-3FN-17373-FokKK_W construct.

### 3.1.3 ZFN expression from ZFN constructs

All ZFNs used in the work presented here were FLAG-tagged. Epitope tagging for antibody binding is a very useful technique to detect the expression (or over-expression) of a specific protein by immunoblotting assay (Western blotting). In order to verify expression of ZFN cassettes constructs, BALB/c 3T3 *scid* m*Tert* fibroblasts were transfected with ZFN plasmids using the calcium phosphate coprecipitation protocol (section 2.2.5) followed by Western blotting analysis as previously described earlier using the ECL method (section 2.2.6.6.1).

Here only plasmids encoding the second and third version of ZFN monomers were used to confirm ZFN expression. The results are shown below in Figure 3.10. Three ZFN plasmid sets were used; the first consisted of ZFNs as pVAX constructs, and the other two sets consisted of ZFNs as pCCL and pRRL-cloned constructs. In all sets there were two different ZFN monomers (ZFN1 or ZFN2) which have different molecular masses; ZFN1 is 42 kDa and ZFN2 is 38.5 kDa. For the two obligate heterodimeric ZFNs in dimeric constructs it is expected to notice 42 and 38.5 kDa proteins, but if they do not separate from each other by N2A self-processing polyprotein signal, the molecular masses would be 80.5 kDa. As expected, Western blot analysis revealed efficient expression from all the constructs in the ZFN pVAX and from most of the ZFN pCCL and pRRL cloned constructs. However, the dimeric constructs encoding for obligatory heterodimeric ZFN1 and ZFN2 together did not show expected expression of 42 and 38.5 kDa proteins, which could be attributed to lack of a N2A self-processing polyprotein signal that helps to separate different protein coding sequences in a single open reading frame (ORF) transcription unit (Tang *et al*., 2009). These findings led to the conclusion that the dimeric constructs were not suitable for further experimentation. All the results were normalised to α-tubulin (49 kDa), detected with a separate antibody and used as a loading control.

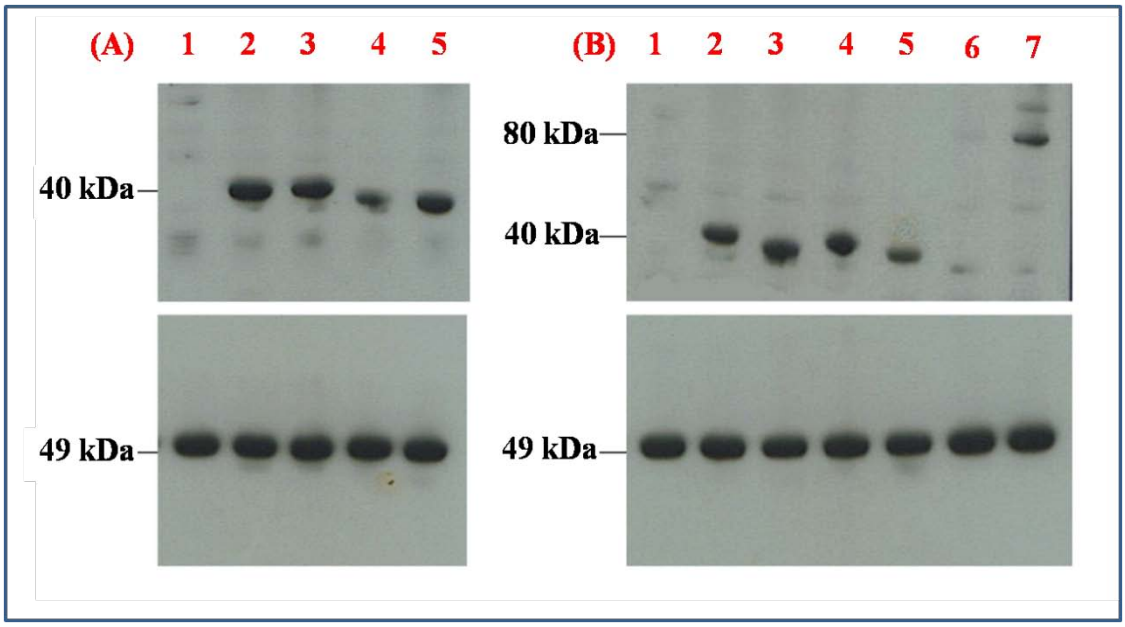



**Figure 3.10: Confirmation of FLAG-tagged ZFN monomer expression from ZFN constructs in BALB/c 3T3 *scid* m*Tert* fibroblasts by Western blotting.** BALB/c 3T3 *scid* m*Tert* fibroblasts were transfected with plasmid constructs by calcium phosphate coprecipitation. Western blotting was completed by ECL protocol. Efficient ZFN expression was detected in all monomeric constructs but not dimeric ones. (A) Expression from original pVAX constructs. Lanes: **1:** Mock, **2:** pVAX-3FN-17373-FokKK (ZFN1 obligate heterodimer), **3:** pVAX-N2A-3FN-17373-Fok (ZFN1 version 2), **4:** pVAX-N2A-3FN-17834-FokEL2 (ZFN2 obligate heterodimer), **5**: pVAX-N2A-3FN-17834-Fok. (ZFN2 version 2). (B) Expression from generated constructs. Lanes: **1**: Mock, **2**: pRRLsc_C_3FN-17373-FokKK_W, **3:** pRRLsc_C_N2A-3FN-17834-FokEL2_W, **4:** pCCLsc_S_3FN-17373-FokKK_W, **5:** pCCLsc_S_N2A-3FN-17834-FokEL2_W, **6:** pRRLsc_C_3FN-17834-FokEL2-3FN-17373-FokKK_W, **7:** pCCLsc_S_3FN-17834-FokEL2-3FN-17373-FokKK_W. Lower panel is detection of mouse α-tubulin (49 kDa) as loading control.

### 3.1.4 Cloning of *Prkdc* short homologous template constructs

The plasmid p*Prkdc*Hind*Bsa*WI was used as a template for PCR amplification of the homologous short *Prkdc*-*Bsa*WI corrective template (of 1626 bp) in order to investigate homologous recombination events in targeted cells as described earlier in section 2.2.12. This shorter template allows for inclusion of a ZFN monomer gene in the same lentiviral backbone. The generated short *Prkdc* template was used to construct new self-inactivating lentiviral vectors encoding for either *Prkdc* template alone or ZFN monomers (driven by either CMV or SFFV promoters) and *Prkdc* template together. The short template was cloned upstream of either promoter and in reverse orientation and contained a diagnostic *Bsa*WI restriction site created by site-directed mutagenesis. This part of work was done in collaboration with Dr. Céline Rocca at RHUL.

#### 3.1.4.1 Construction of pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc*

This plasmid was constructed to encode both CMV driven-ZFN1 and short *Prkdc* corrective template from the same lentiviral construct. The genetic maps of the recipient CMV driven-ZFN1 plasmid (pRRLsc_C_3FN-17373-FokKK_W), plasmid used as a PCR template (p*Prkdc*Hind*Bsa*WI), and the new construct (pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc*) are all shown in Figure 3.11. The construction of lentivector transfer plasmid expressing ZFN1 heterodimer driven by CMV promoter and *Prkdc* short template was completed as follows: PCR amplification was done as described in (section 2.2.12) using p*Prkdc*Hind*Bsa*WI generating a *Prkdc* short template which was then digested with *San*DI. Next, the PCR amplicon of 1626 bp was cloned into pRRLsc_C_3FN-17373-FokKK_W previously digested with *San*DI. A plasmid clone in which the *Prkdc* template had inserted in the reverse orientation was selected. The generated plasmid 9514 bp was verified by characteristic restriction digest using *Spe*I, generating a fragment of 9514 bp, and *Eco*RI generating fragments consisting of 5517 and 3997 bp; whereas *Nco*I generated fragments of 6850, 1911 and 753 bp.

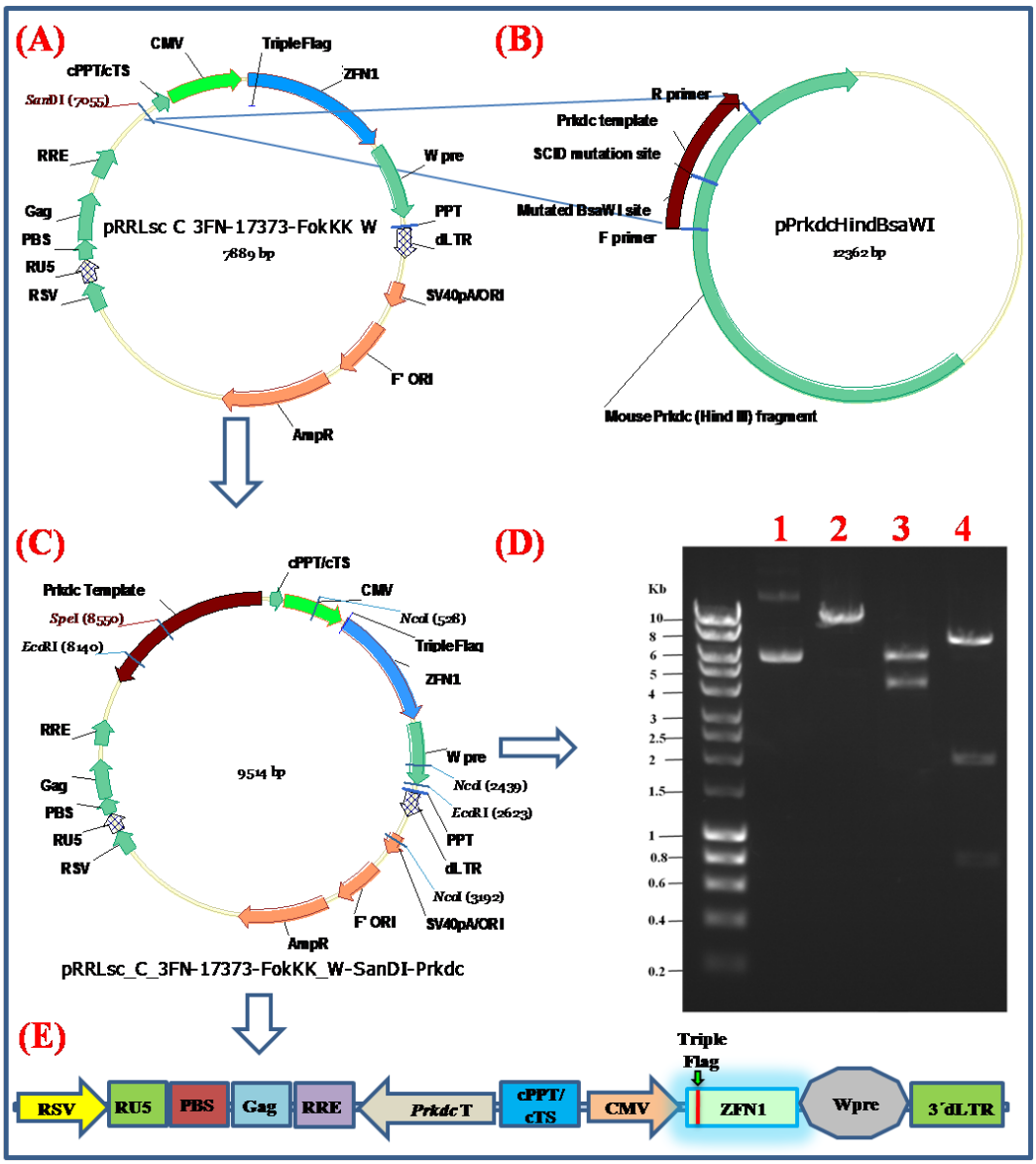



**Figure 3.11: Construction and characterisation of pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc*.** (A), (B) and (C) show genetic maps of (pRRLsc_C_3FN-17373-FokKK_W), (p*Prkdc*Hind*Bsa*WI) and (pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc*); respectively. (D) Restriction endonuclease analysis of generated pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc* (9514 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Spe*I generating fragment size of 9514 bp, lane 3 digestion with *Eco*RI generating fragment sizes of 5517 and 3997 bp, lane 4 digestion with *Nco*I generating fragment sizes of 6850, 1911 and 753. DNA ladder: 5 µl of Bioline Hyperladder I. (E) Depicts a general schematic diagram of pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc* construct.

### *3.1.4.2* Construction of pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*

This plasmid was constructed to encode both the CMV driven ZFN2 and the short *Prkdc* corrective template in the same lentiviral construct. The genetic maps of recipient CMV driven-ZFN2 plasmid (pRRLsc_C_N2A-3FN-17834-FokEL2_W), the plasmid used as a PCR template (p*Prkdc*Hind*Bsa*WI), and the new construct (pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*) are all shown in Figure 3.12. The construction of the lentiviral transfer plasmid expressing ZFN2 heterodimer driven by CMV promoter and *Prkdc* short template was completed as follows: PCR amplification was done as described in section 2.2.12 using p*Prkdc*Hind*Bsa*WI as a template, generating *Prkdc* short template which was then digested with *San*DI. Next, the PCR amplicon of 1626 bp was cloned into pRRLsc_C_N2A-3FN-17834-FokEL2_W which has been previously digested with *San*DI. A plasmid clone in which the *Prkdc* template had inserted in the reverse orientation was selected. The generated plasmid (9415 bp) was verified by characteristic restriction using *Spe*I generating fragment size of 9415 bp, *Eco*RI generating fragment sizes of 5517 and 3898 bp and *Nco*I generating fragment sizes of 6850, 1501, 753 and 311 bp.

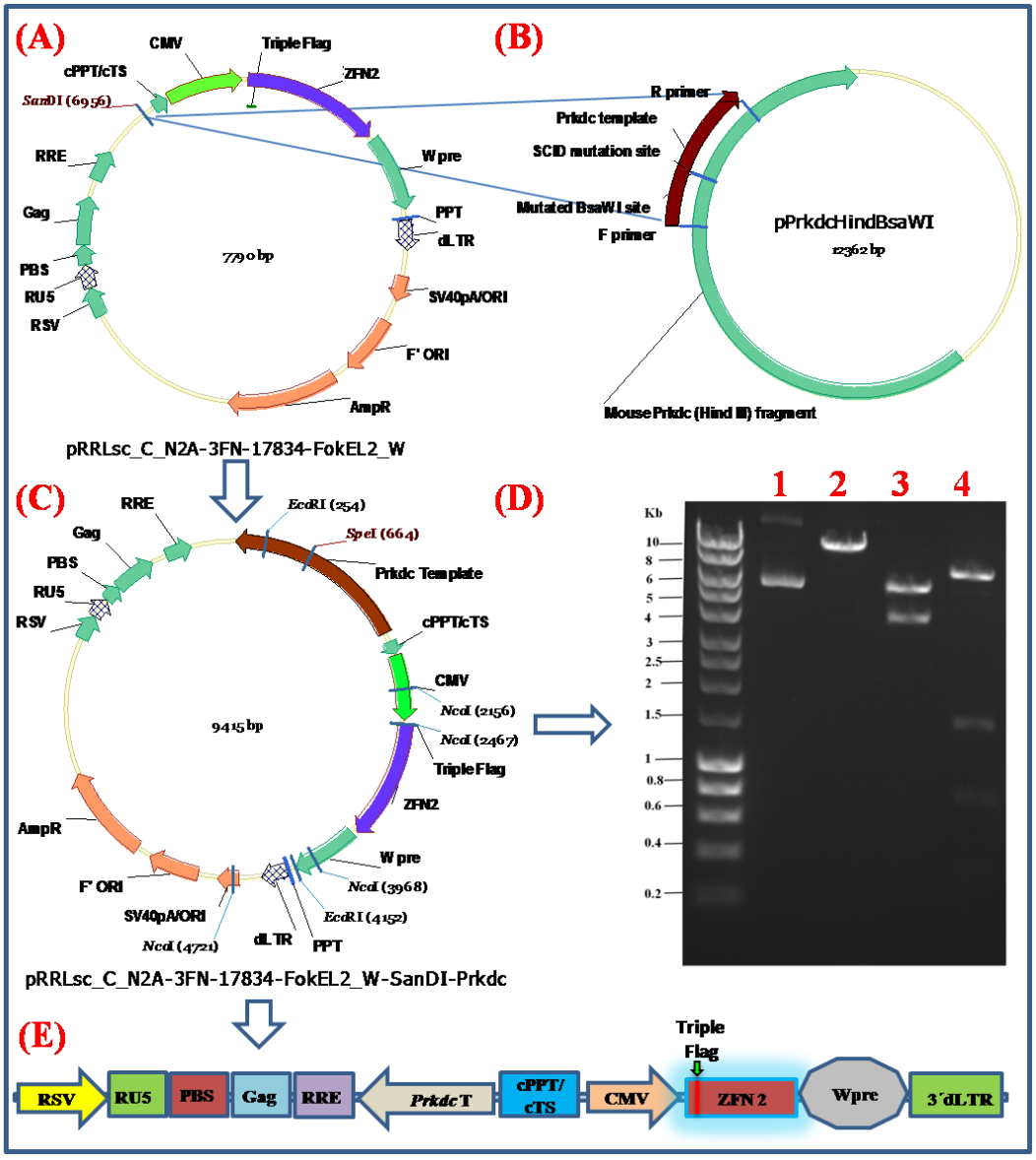



**Figure 3.12: Construction and characterisation of pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*.** (A), (B) and (C) show genetic maps of (pRRLsc_C_N2A-3FN-17834-FokEL2_W), (p*Prkdc*Hind*Bsa*WI) and (pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*); respectively. (D) Restriction endonuclease analysis of generated pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* (9415 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Spe*I generating fragment size of 9415 bp, lane 3 digestion with *Eco*RI generating fragment sizes of 5517 and 3898 bp, lane 4 digestion with *Nco*I generating fragment sizes of 6850, 1501, 753 and 311. DNA ladder: 5 µl of Bioline Hyperladder I. (E) Depicts a general schematic of pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* construct.

### *3.1.4.3* Construction of pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc*

Here the aim was to clone both SFFV driven-ZFN1 heterodimer and short *Prkdc* corrective template into the same lentiviral construct. The genetic maps of the recipient SFFV-driven ZFN1 plasmid (pCCLsc_S_3FN-17373-FokKK_W), the plasmid used as a PCR template (p*Prkdc*Hind*Bsa*WI), and the new construct (pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc*) are all shown in Figure 3**.**13. The construction of lentiviral transfer plasmid expressing ZFN1 heterodimer driven by a SFFV promoter and *Prkdc* short template was completed as follows: PCR amplification was carried out as described in section 2.2.12.1 using p*Prkdc*Hind*Bsa*WI generating *Prkdc* short template which was then digested with *San*DI. Next, the PCR amplicon of 1626 bp was cloned into pCCLsc_S_3FN-17373-FokKK_W previously digested with *San*DI. A plasmid clone in which the *Prkdc* template had inserted in the reverse orientation was selected. The generated plasmid (9872 bp) was verified by characteristic restriction using *Not*I generating fragment size of 9872 bp, *Spe*I generating fragment sizes of 7151 and 2721 bp, *Nco*I generating fragment sizes of 5678, 3441 and 753 bp and *Eco*RI generating fragment sizes of 5960, 1791, 1598 and 523 bp.

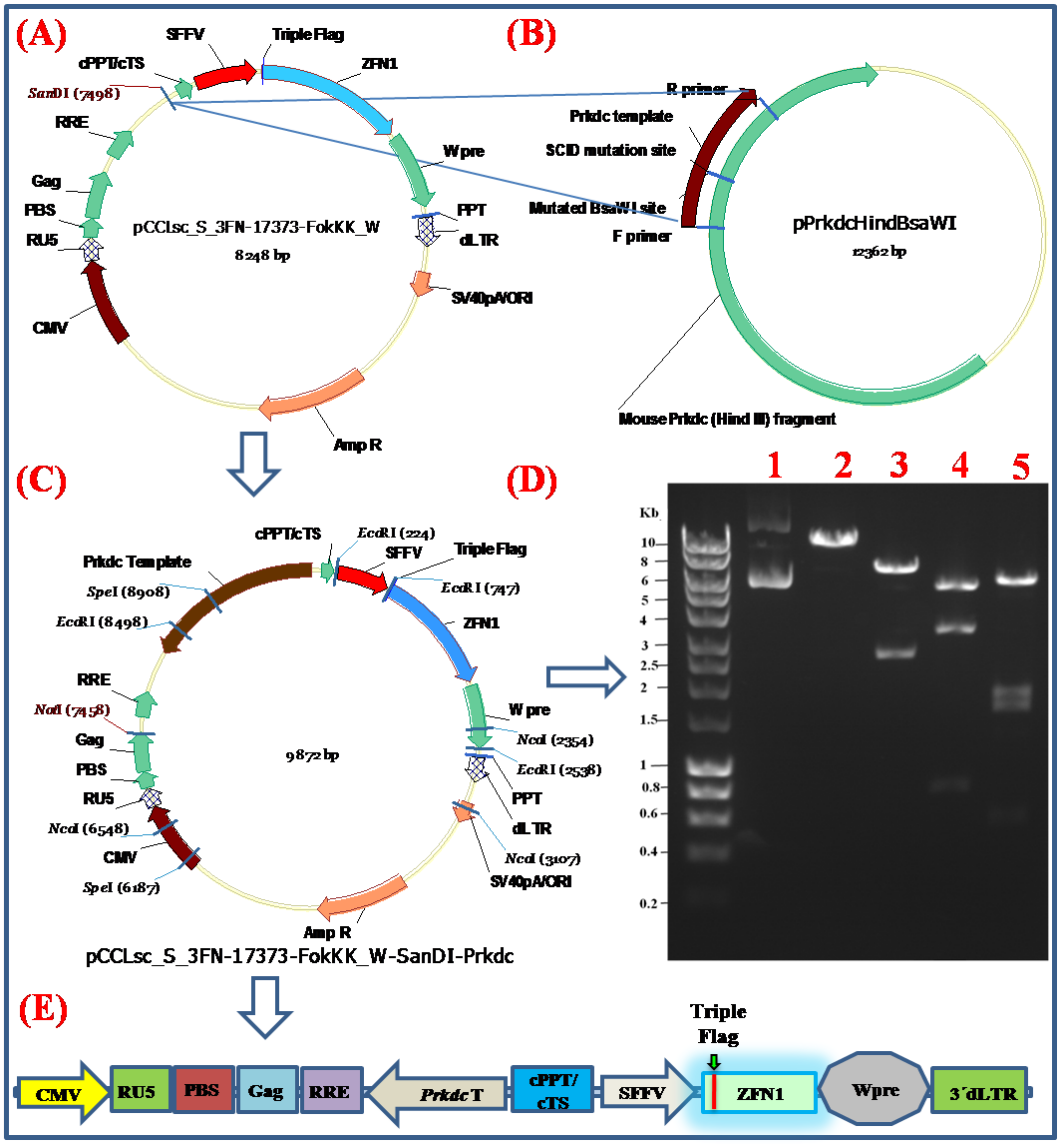



**Figure 3.13: Construction and characterisation of pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc*.** (A), (B) and (C) show genetic maps of (pCCLsc_S_3FN-17373-FokKK_W), (p*Prkdc*Hind*Bsa*WI) and (pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc*); respectively. (D) Restriction endonuclease analysis of generated pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc* (9872 bp) construct on 0.7% agarose gel. Lane 1 undigested product, lane 2 digestion with *Not*I generating fragment size of (9872) bp, lane 3 digestion with *Spe*I generating fragment sizes of 7151 and 2721 bp, lane 4 digestion with *Nco*I generating fragment sizes of 5678, 3441 and 753 bp and lane 5 digestion with *Eco*RI generating fragment sizes of 5960, 1791, 1598 and 523 bp.DNA ladder: 5 µl of Bioline Hyperladder I. (E) Depicts a schematic diagram of the pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc* construct.

### *3.1.4.4* Construction of pCCLsc_S _N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*

This plasmid constructed to encode both the SFFV-driven ZFN2 heterodimeric monomer and the short *Prkdc* corrective template in the same lentiviral construct. The genetic maps of the recipient SFFV-driven ZFN2 plasmid (pCCLsc_S_N2A-3FN-17834-FokEL2_W), plasmid used as a PCR template (p*Prkdc*Hind*Bsa*WI), and the new construct (pCCLsc_S _N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*) are shown in Figure 3.14. The construction of lentiviral transfer plasmid expressing ZFN2 heterodimer driven by the SFFV promoter and *Prkdc* short template was completed as follows: PCR amplification was done as described in section 2.2.12 using p*Prkdc*Hind*Bsa*WI, generating *Prkdc* short template which was then digested with *San*DI. Next, the PCR amplicon of 1626 bp was ligated into pCCLsc_S_N2A-3FN-17834-FokEL2_W previously digested with *San*DI. A plasmid clone in which the *Prkdc* template had inserted in the reverse orientation was selected. The generated plasmid (9774 bp) was verified by characteristic restriction using *Not*I generating fragment size of 9774 bp, *Spe*I generating fragment sizes of 7053 and 2721 bp, *Nco*I generating fragment sizes of 4079, 3441, 1501 and 753 bp and *EcoR*I generating fragment sizes of 4079, 3441, 1501 and 753 bp.

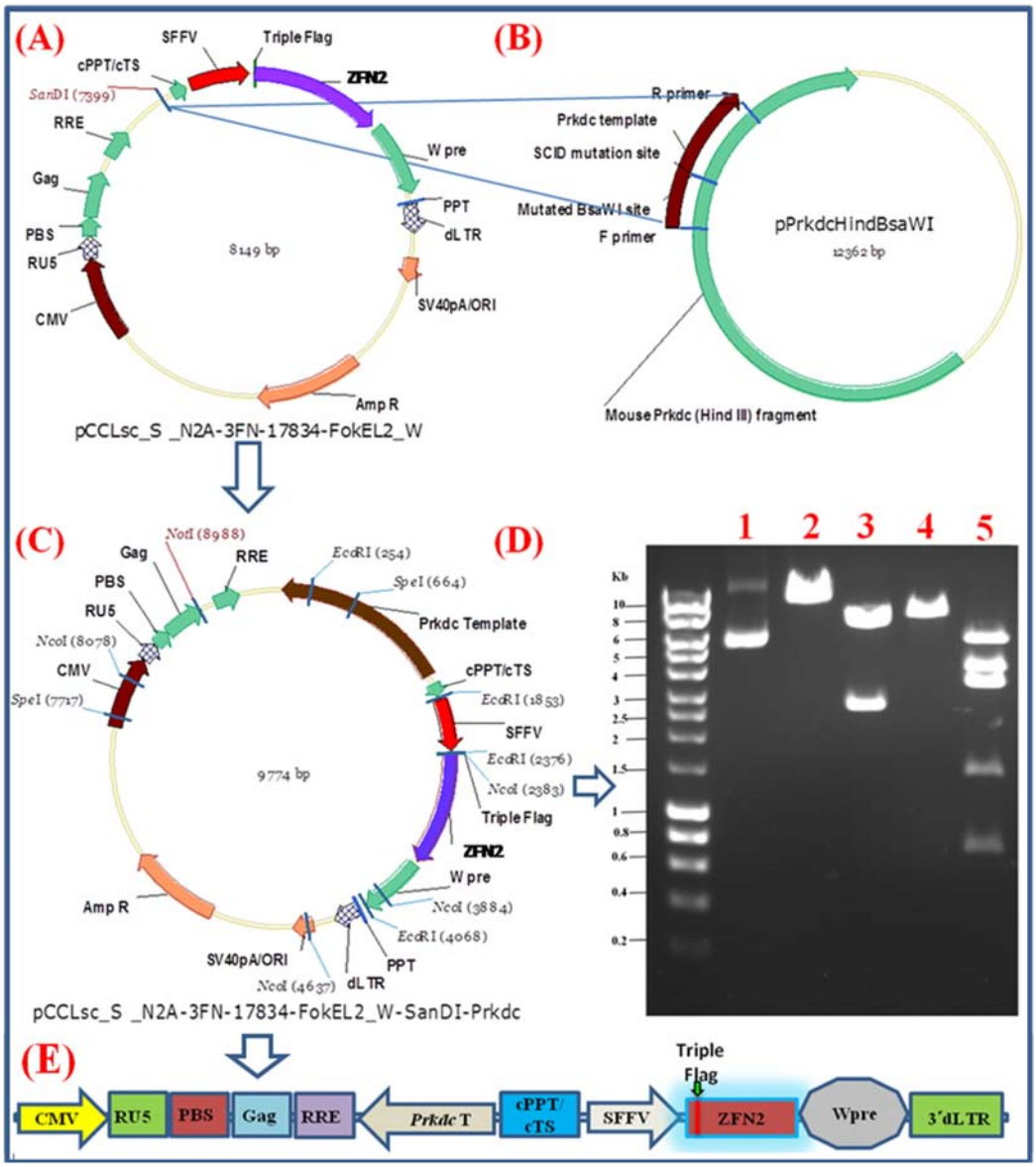



**Figure 3.14: Construction and characterisation pCCLsc_S _N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*.** (A), (B) and (C) show genetic maps of (pRRLsc_C_N2A-3FN-17834-FokEL2_W), (p*Prkdc*Hind*Bsa*WI) and (pCCLsc_S_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*); respectively. (D) Restriction endonuclease analysis of generated pCCLsc_S _N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* (9774 bp) construct on 0.7% agarose gel. Lane 1: undigested product, lane 2 digestion with *Not*I generating fragment size of 9774 bp, lane 3 digestion with *Spe*I generating fragment sizes of 7053 and 2721 bp, lane 4 digestion with *Nco*I generating fragment sizes of 4079, 3441, 1501 and 753 bp and lane 5 digestion with *Eco*RI generating fragment sizes of 4079, 3441, 1501 and 753 bp. DNA ladder: 5 µl of Bioline Hyperladder I. (E) A schematic of pCCLsc_S _N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* construct.

### *3.1.4.5* Construction of pCCLsc-*San*DI-*Prkdc*

This construct was designed to only contain the *Prkdc* template. The genetic maps of the donor plasmid (pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc*) and the new construct (pCCLsc-*San*DI-*Prkdc*) are shown in Figure 3.15. The construction of the lentiviral transfer plasmid expressing only the *Prkdc* short template was completed as follows: the construct pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc* was double digested with *Eco*RV and *Kpn*I, then the largest fragment (7538 bp) was purified and self-ligated, generating the pCCLsc-*San*DI-*Prkdc* construct. The generated plasmid 7539 bp was verified by characteristic restriction using *Nco*I generating fragment sizes of 4098 and 3441 bp and *Spe*I generating fragment sizes of 4818 and 2721 bp.

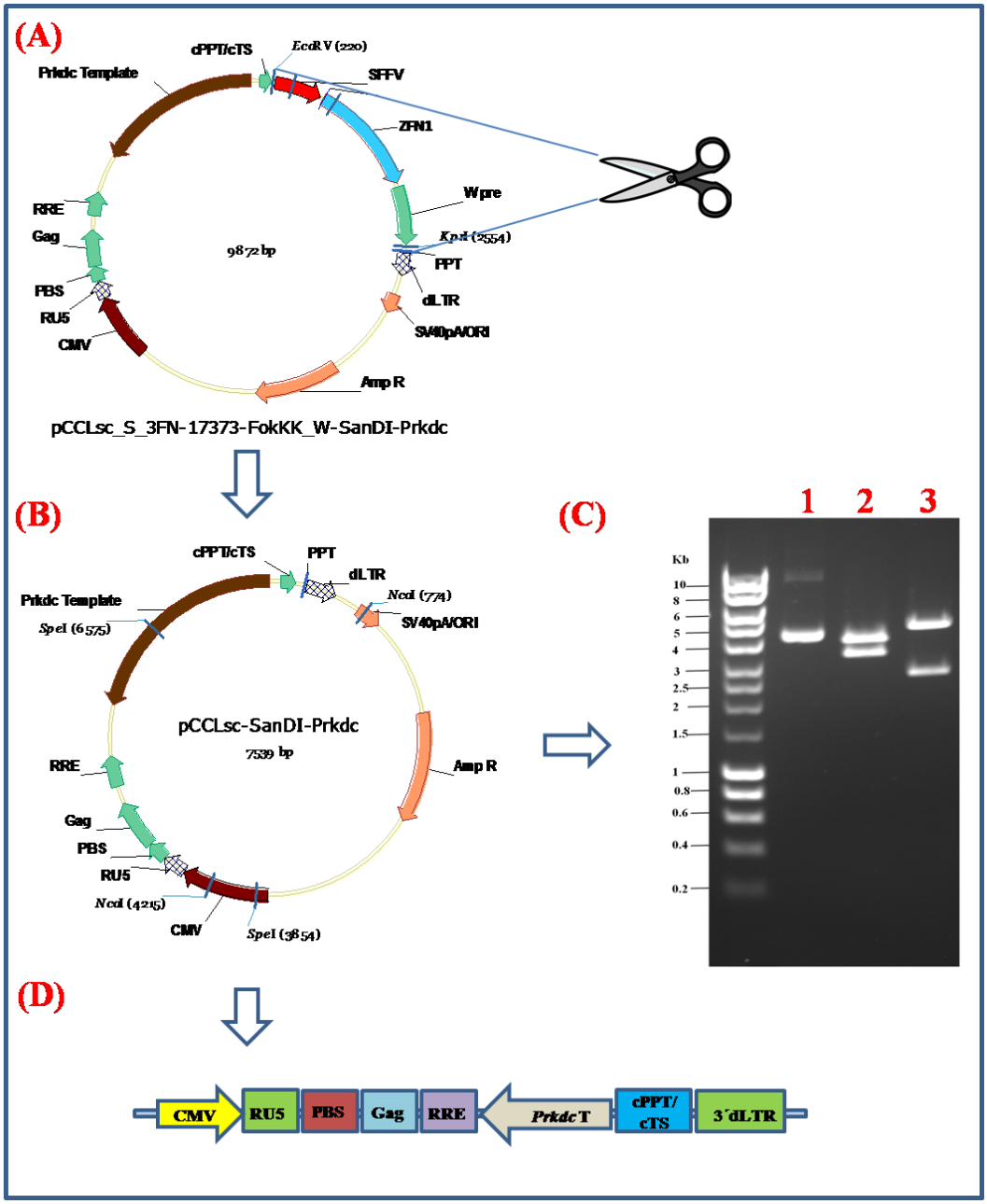



**Figure 3.15: Construction and characterisation of pCCLsc-*San*DI-*Prkdc*.** (A) and (B) show genetic maps of (pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc*) and (pCCLsc-*San*DI-*Prkdc*); respectively. (D) Restriction endonuclease analysis of generated pCCLsc-*San*DI-*Prkdc* (7539 bp) construct on 0.7% agarose gel. Lane 1: undigested product, lane 2: digestion with *Nco*I generating fragment sizes of 4098 and 3441 bp and lane 3: digestion with *Spe*I generating fragment sizes of 4818 and 2721 bp. DNA ladder: 5 µl of Bioline Hyperladder I. (E) A schematic of the pCCLsc-*San*DI-*Prkdc* construct.

### 3.1.5 ZFN expression from ZFN-*Prkdc* short homologous template constructs

Lentiviral ZFN constructs that were generated in section 3.1.2 , were used again to clone the wt short *Prkdc* template upstream of both CMV & SFFV promoters and in reverse orientation (see section 3.1.4). Importantly all the original features were kept intact, including the FLAG-tag. Aiming to detect ZFN expression from ZFN-*Prkdc* short homologous template constructs, BALB/c 3T3 *scid* m*Tert* fibroblasts were transfected with ZFN plasmids using the calcium phosphate coprecipitation method (section 2.2.5) followed by Western blotting analysis as previously described (section 2.2.6.6.2). The results are shown below in Figure 3.16. Efficient expression from all constructs encoding for ZFN1 and ZFN2 (42 and 38.5 kDa, respectively) was observed. Mouse α-tubulin (49 kDa) was used as loading control.

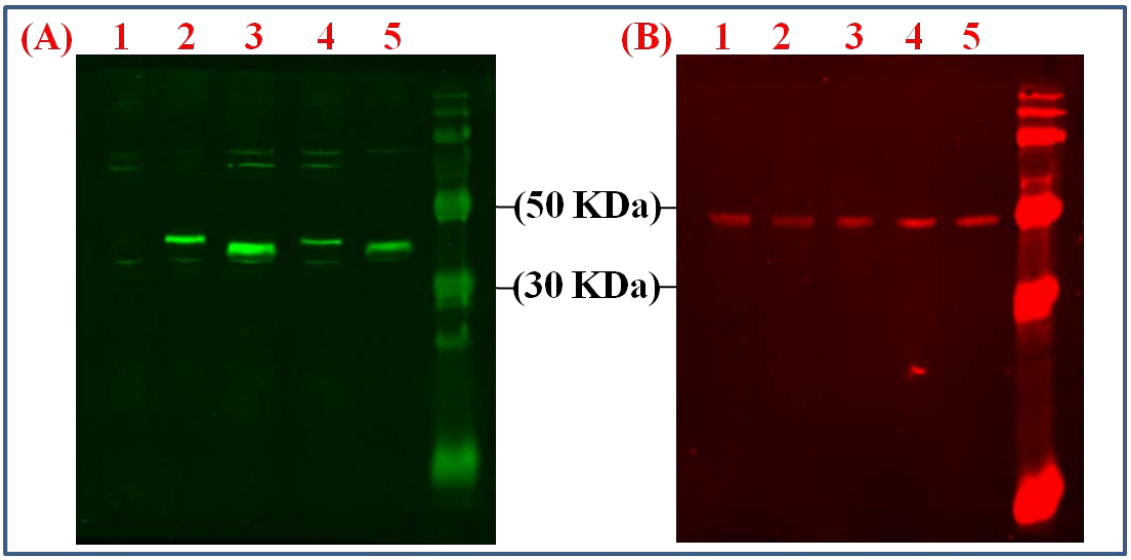



**Figure 3.16: Confirmation of FLAG-tagged ZFN expression from ZFN-*Prkdc* short homologous template constructs in BALB/c 3T3 *scid* m*Tert* fibroblasts by Western blotting.** BALB/c 3T3 *scid* m*Tert* fibroblasts were transfected with ZFN-*Prkdc* short homologous template constructs by calcium phosphate coprecipitation. Western blotting was completed by the Odyssey infrared imaging method. Efficient ZFN expression was detected from all constructs; mouse α-tubulin was used as a loading control. (A) Expression from ZFN-*Prkdc* short homologous template constructs. Lanes: 1: Mock, 2:pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc*, 3: pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*, 4: pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc*, 5: pCCLsc_S_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc*. (B) Detection of mouse α-tubulin (49 kDa) as loading control.

## 3.2 Lentiviral vector production

The viral enzyme integrase plays a vital role in lentiviral integration. As I have described in the Introduction, mutations within the viral integrase gene have been identified which allow for production of two different types of lentiviral vectors: IPLVs which encode an active integrase, whereas IDLVs encode a catalytically inactive viral integrase due to a point mutation in the integrase coding region, creating a D64V change in the amino acid sequence (Yanez-Munoz *et al*., 2006). IDLVs are becoming increasingly promising in different gene therapy applications as they are transcriptionally active although at lower levels than the IPLVs (Negri *et al*., 2010). In addition, they also show efficient transgene expression (Chick *et al*., 2012), offering a reduced risk of insertional mutagenesis within targeted genomes due to their episomal nature (Wanisch and Yanez-Munoz, 2009).

In current study, all lentiviral vectors were produced by calcium phosphate transient co-transfection of HEK 293T cells. Table 3.2 lists all lentiviral vectors prepared throughout current study with their main features. All vectors were pseudotyped with the VSV-G envelop using the plasmid pMD2.VSV-G to express the VSV-G from the human CMV immediate early promoter as previously described by (Naldini *et al*., 1996). Lentiviral vectors encoding *Prkdc* templates were only made in IDLV configuration as the genomic integration of the template should be avoided. All other vectors were made as IPLVs and IDLVs. ZFN-encoding lentiviral vectors produced obligatory heterodimers. Lentiviral vectors with eGFP cassettes were either driven by the CMV promoter (using transfer plasmid pRRLsc-CEW) or by the SFFV promoter (using transfer plasmid pHR'sc_SEW).

Real time quantitative PCR (qRT-PCR) was used in this study to titrate all lentiviral vectors, as it allows for a simultaneous amplification and quantification of target DNA using a fluorescently labelled reporter. There are three possible configurations for lentiviral vector cDNA in transduced cells: linear double stranded DNA, chromosomally integrated provirus, or extrachromosomal circular products with 1 LTR or 2 LTRs. Real time qPCR titration was conducted as described in section 2.2.7.3. The HIV LRT reaction can measure all HIV cDNA varieties (Yoder and Fishel, 2008). All

prepared lentiviral vectors, including eGFP, were titrated by real time qPCR and their titre range was $2 \times 10^9$-$6.0 \times 10^{10}$.

On the other hand, lentiviral vectors expressing eGFP were titrated by flow cytometry in addition to their titration by qPCR; flow cytometry titration was done as described in section 2.2.7.2 and based on preparation of serially diluted lentiviral vectors transduced into HeLa cells. As an example of eGFP titre calculation, Figure 3.17 shows a representative flow cytometric result. A series of ten-fold dilutions of the original viral suspension was used to transduce test cells. Those samples that gave a range of (1-10 %) eGFP positive cells were used to determine the titre. The flow cytometry (eGFP) titre range for the lentiviral vectors used in this Thesis was $4 \times 10^8 - 2 \times 10^9$.

**Table 3.2: Main features of lentiviral vectors prepared throughout current study.**

| ≠ | Transfer plasmid | Transgene | Integration efficiency | Transgene promoter | Figure ≠ |
|---|---|---|---|---|---|
| 1 | pRRLsc-CEW | eGFP cassette | IPLV | CMV | N/A |
| 2 | pRRLsc-CEW | eGFP cassette | IDLV | CMV | N/A |
| 3 | pHR'sc-SEW | eGFP cassette | IPLV | SFFV | N/A |
| 4 | pHR'sc-SEW | eGFP cassette | IDLV | SFFV | N/A |
| 5 | pCCLsc_*Prkdc*Hind*Bsa*WIF_ISceIT | *Prkdc-Bsa*WI template (7389 bp) | IDLV | CMV | 3.1 |
| 6 | pCCLsc_*Prkdc*Hind*Bsa*WIR_ISceIT | *Prkdc-Bsa*WI template (7389 bp) | IDLV | CMV | 3.2 |
| 7 | pCCLsc-*San*DI-*Prkdc* | *Prkdc-Bsa*WI template (1626 bp) | IDLV | CMV | 3.15 |
| 8 | pRRLsc_C_3FN-17373-FokKK_W | ZFN1 hetrodimer-V3 | IPLV | CMV | 3.3 |
| 9 | pRRLsc_C_N2A-3FN-17834-FokEL2_W | ZFN2 hetrodimer-V3 | IPLV | CMV | 3.4 |
| 10 | pCCLsc_S_3FN-17373-FokKK_W | ZFN1 hetrodimer-V3 | IPLV | SFFV | 3.5 |
| 11 | pCCLsc_S _N2A-3FN-17834-FokEL2_W | ZFN2 hetrodimer-V3 | IPLV | SFFV | 3.6 |
| 12 | pRRLsc_C_3FN-17373-FokKK_W | ZFN1 hetrodimer-V3 | IDLV | CMV | 3.3 |
| 13 | pRRLsc_C_N2A-3FN-17834-FokEL2_W | ZFN2hetrodimer-V3 | IDLV | CMV | 3.4 |
| 14 | pCCLsc_S_3FN-17373-FokKK_W | ZFN1 hetrodimer-V3 | IDLV | SFFV | 3.5 |
| 15 | pCCLsc_S _N2A-3FN-17834-FokEL2_W | ZFN2 hetrodimer-V3 | IDLV | SFFV | 3.6 |
| 3 | pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc* | ZFN1 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IPLV | CMV | 3.11 |
| 17 | pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* | ZFN2 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IPLV | CMV | 3.12 |
| 18 | pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc* | ZFN1 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IPLV | SFFV | 3.13 |
| 19 | pCCLsc_S _N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* | ZFN2 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IPLV | SFFV | 3.14 |
| 20 | pRRLsc_C_3FN-17373-FokKK_W-*San*DI-*Prkdc* | ZFN1 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IDLV | CMV | 3.11 |
| 21 | pRRLsc_C_N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* | ZFN2 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IDLV | CMV | 3.12 |
| 22 | pCCLsc_S_3FN-17373-FokKK_W-*San*DI-*Prkdc* | ZFN1 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IDLV | SFFV | 3.13 |
| 23 | pCCLsc_S _N2A-3FN-17834-FokEL2_W-*San*DI-*Prkdc* | ZFN2 hetrodimer-V3 & *Prkdc-Bsa*WI template (1626 bp) | IDLV | SFFV | 3.14 |

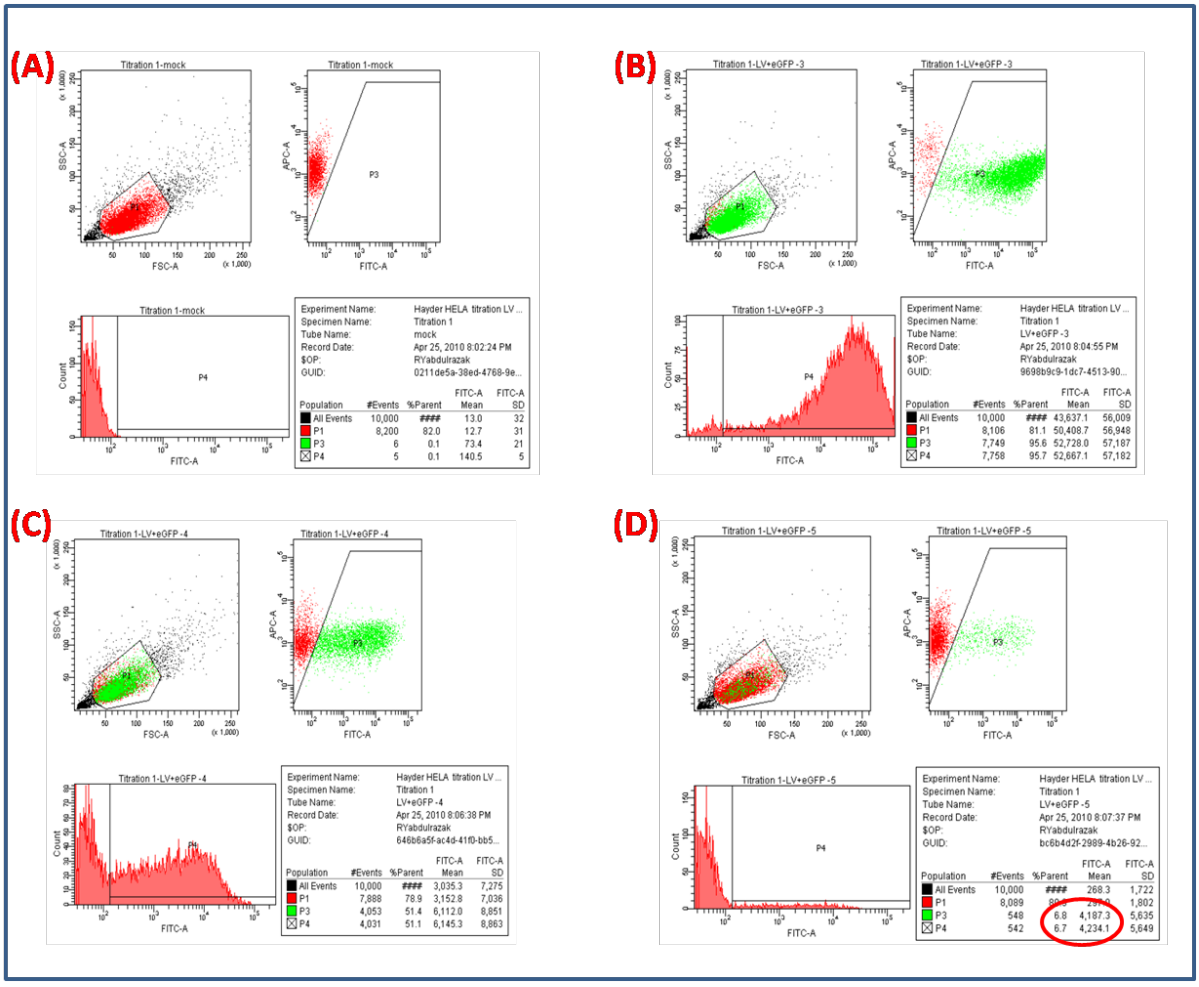



**Figure 3.17: Representative flow cytometry results for eGFP titration.** HeLa cells were transduced with serial ten-fold dilutions of eGFP lentiviral vectors**,** harvested three days later and analysed by flow cytometry. The FITC-A channel was used to detect eGFP fluorescence, and the APC-A channel for cell auto-fluorescence. Dilutions that gave 1-10 % eGFP -positive cells were used to calculate the titre using the formula: eGFP transducing units (TU)/ml=% green cells x$10^5$ (cells/well at the time of vector addition) x 1/vector dilution. (A) Mock sample, (B, C&D) are tenfold dilution samples. Panel (D) represents the percentage of green cells that will be used in eGFP titre calculation for this vector (region P3).

### 3.2.1 ZFN expression from ZFN lentiviral vectors

To test the level of expression of ZFN monomers expressed from lentiviral vectors, ZFN IPLVs and IDLVs encoding FLAG-tagged ZFN monomers driven from either CMV or SFFV promoters were used to transduce BALB/c 3T3 *scid* m*Tert* fibroblasts. The MOI of 100 (qPCR) was used in the presence of polybrene (8 µg/ml). The presence of the FLAG-tagged ZFN protein was then confirmed by Western blotting using Odyssey infrared imaging method. Results are shown below in Figures 3.18, 3.19 & 3.20. The expected molecular mass of ZFN1 is 42 kDa and of ZFN2 is 38.5 kDa. Mouse α-tubulin (49 kDa) was used as loading control. Western blot analysis revealed a strong expression from ZFN IPLVs, while lower expression (particularly from SFFV promoter) was observed with ZFN IDLVs, as expected for transduction of dividing cells. There was no noticeable difference between expression level from CMV or SFFV promoters in either the IPLV or IDLV ZFNs configuration.

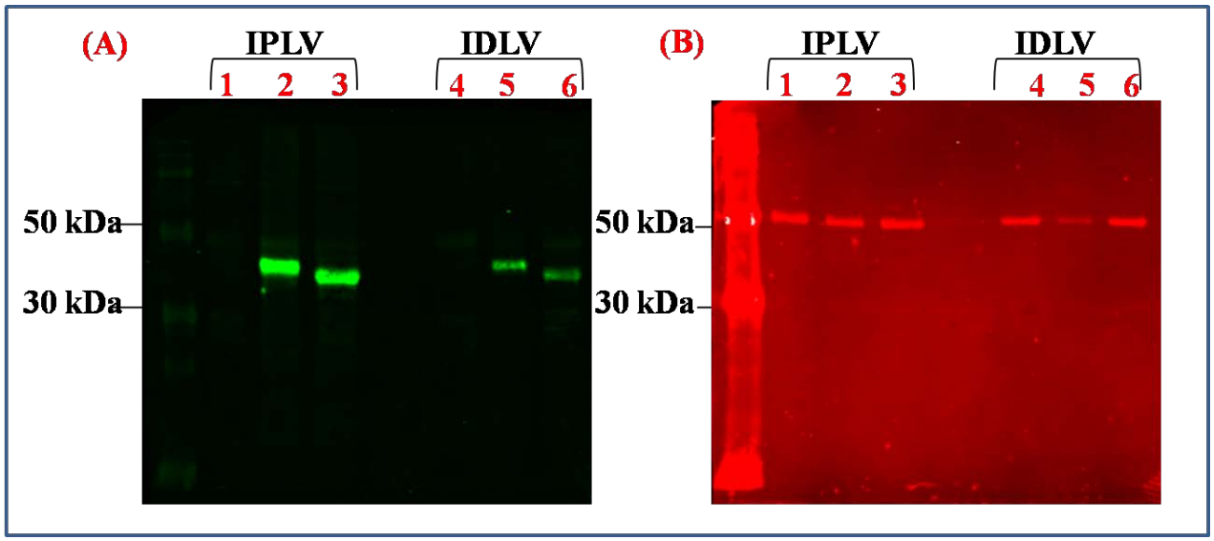



**Figure 3.18: Confirmation of FLAG-tagged ZFN expression from CMV-ZFN lentiviral vectors in BALB/c 3T3 *scid* m*Tert* fibroblasts by Western blotting.** BALB/c 3T3 *scid* m*Tert* fibroblasts were transduced with CMV ZFN IPLVs or IDLVs fibroblasts with qPCR MOI 100 in presence of polybrene (8 µg/ml). Western blotting was performed by the Odyssey infrared imaging method. Efficient ZFN expression was detected from ZFN IPLVs, whereas lower expression was observed from ZFN IDLVs. (A) FLAG-tagged ZFN expression from IPLVs and IDLVs. Lanes: 1& 4: Mock, 2& 5: ZFN1, 3 &6: ZFN2. (B) Detection of mouse α-tubulin (49 kDa) as loading control.

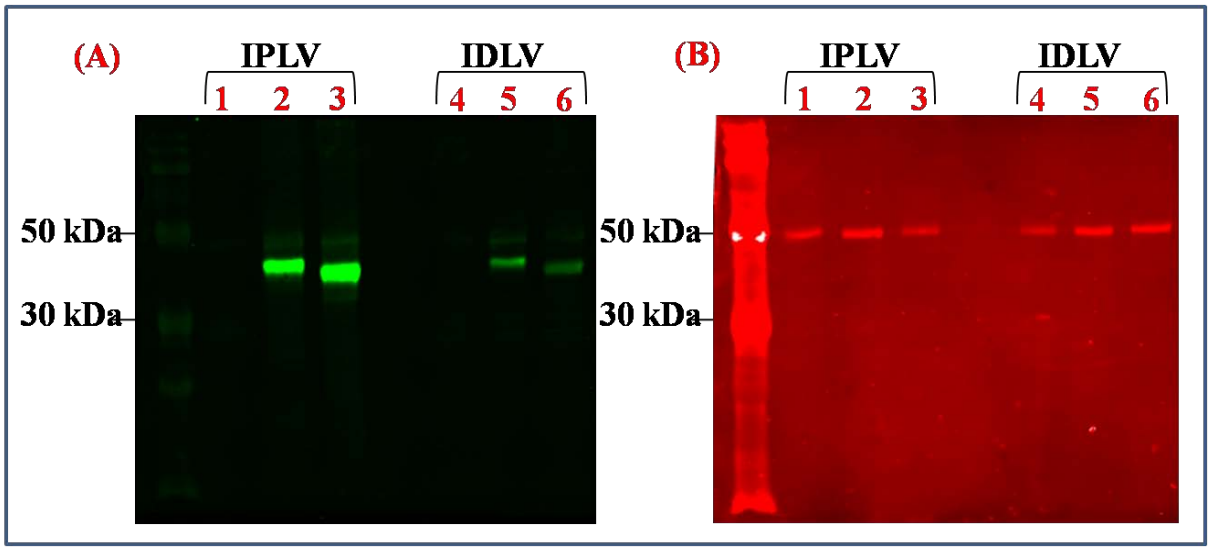



**Figure 3.19: Confirmation of FLAG-tagged ZFN expression from SFFV-ZFN lentiviral vectors in BALB/c 3T3 *scid* m*Tert* fibroblasts by Western blotting.** BALB/c 3T3 *scid* m*Tert* fibroblasts were transduced with SFFV ZFN IPLVs or IDLVs fibroblasts with qPCR MOI 100 in the presence of polybrene (8 µg/ml). Western blotting was performed by the Odyssey infrared imaging method. Efficient ZFN expression was detected from ZFN IPLVs, whereas significantly lower expression was observed from ZFN IDLVs. (A) FLAG-tagged ZFN expression from IPLVs and IDLVs. Lanes: 1& 4: Mock, 2& 5: ZFN1, 3 &6: ZFN2. (B) Detection of mouse α-tubulin (49 kDa) as loading control.

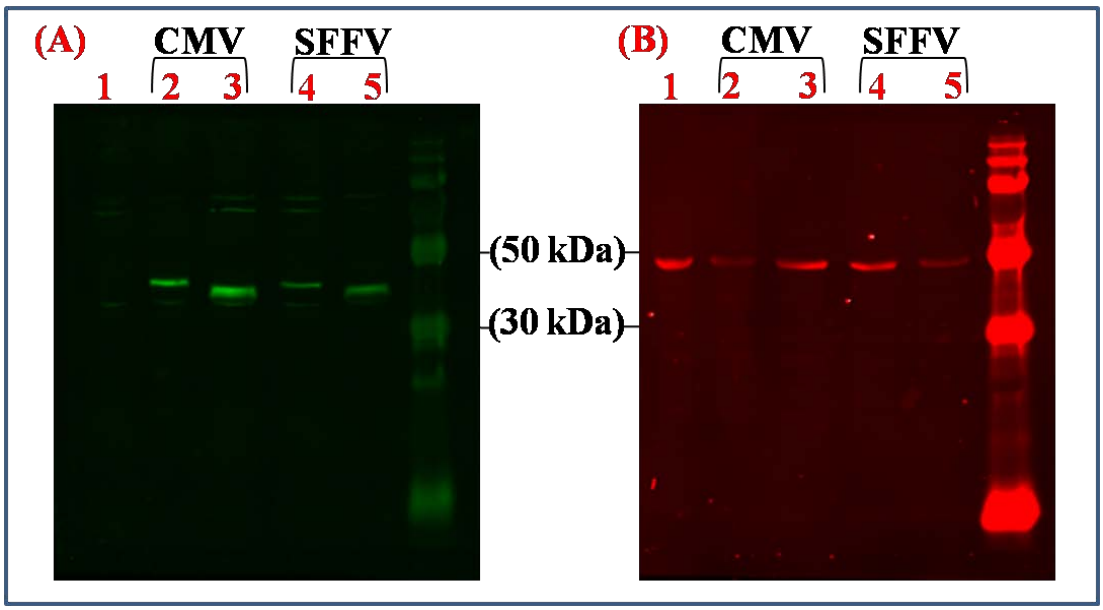



**Figure 3.20: Confirmation of FLAG-tagged ZFN expression from ZFN-*Prkdc* short template IDLVs in BALB/c 3T3 *scid* m*Tert* fibroblasts by Western blotting.** BALB/c 3T3 *scid* m*Tert* fibroblasts were transduced with ZFN-*Prkdc* short template IDLVs with qPCR MOI 100 in presence of polybrene (8 µg/ml). Western blotting was performed by the Odyssey infrared imaging method. ZFN expression was detected from all ZFN IDLVs. (A) FLAG-tagged ZFN expression from IDLVs. Lanes: 1: Mock, 2: *Prkdc*/CMV ZFN1, 3: *Prkdc*/CMV ZFN2, 4: *Prkdc*/SFFV ZFN1, 5: *Prkdc*/SFFV ZFN2. (B) Detection of mouse α-tubulin (49 kDa) as loading control.

# Chapter Four

# ZFN activity in *scid* fibroblasts

# 4 ZFN activity in *scid* fibroblasts

This chapter can be divided into two parts, in the first part I describe three preliminary experiments performed with the aim of generating tools or optimising procedures to be used in the context of the current study. First, BALB/c 3T3 *scid* fibroblasts were immortalised to overcome their self senescence which was followed by assessing the effect of polybrene on transduction efficiency of BALB/c 3T3 *scid* m*Tert* fibroblasts, and optimisation of plasmid transfection by calcium phosphate coprecipitation method. The second and main part of the chapter presents data to use the tools generated in previous experiments to test the ZFNs *in vitro* by using BALB/c 3T3 *scid* m*Tert* fibroblasts. The vital aims were to test how efficiently the ZFNs would recognize their targets, to detect gene repair events and to demonstrate restoration of *Prkdc* functionality. In order to do so, first of all I examined different ZFN variants using a gene targeting assay to select ZFNs that could induce higher levels of DSB-mediated HR. Then I analysed gene targeting results with PCR assay and subsequent digestion with restriction enzymes to confirm HR events. Later, I repeated gene targeting experiment but with ZFN IPLVs this time to check out ZFN cassettes expression from prepared ZFN lentiviral vectors. Next I worked on demonstration of efficiency of ZFNs cutting by surveyor's *Cel*-I assay. I have done another gene repair experiment and examined HR events by *Bsa*WI assay. Subsequently, in order to check restoration of functionality of *Prkdc*; DNA-PK activity assay was performed. Moreover; to further confirm restoration of *Prkdc* activity, first, melphalan drug cytotoxic dose for BALB/c 3T3 wt and *scid* fibroblasts was detected by MTT assay. Then corrected cells were enriched by culturing them in presence of melphalan optimised dose and then their growth was analysed. Finally *Bsa*WI assay was repeated to compare between gene repair levels of corrected cells before and after enrichment with melphalan.

## 4.1 Immortalisation of BALB/c 3T3 *scid* fibroblasts

BALB/c 3T3 *scid* fibroblasts were used in the current study as convenient target cells to demonstrate ZFN cutting and gene repair. Work performed prior to my arrival in the laboratory had demonstrated that these cells suffered from extensive senescence upon cloning. With the intention of extending their life-span and overcoming senescence, BALB/c 3T3 *scid* fibroblasts were immortalised by transduction with retroviral vectors encoding mouse telomerase reverse transcriptase *(Tert)*, the gene encoding for the catalytic subunit of TERT. Immortalised fibroblasts are referred to as BALB/c 3T3 *scid* m*Tert* fibroblasts and they were widely used in this study.

### 4.1.1 Titration of puromycin resistance of mouse *scid* fibroblasts

As mentioned above, initial work carried out using BALB/c 3T3 *scid* fibroblasts was hindered by their extensive senescence. The retroviral vector used for the immortalisation encoded both the puromycin resistance gene *Pac* and m*Tert*, the former allowing selection of transduced cells. Puromycin resistance is conferred by the acquisition and expression of the *Pac* gene, encoding a puromycin N-acetyl transferase (Sanchez-Puig and Blasco, 2000). It was important to titrate puromycin susceptibility in those fibroblasts before immortalisation because BALB/c 3T3 *scid* m*Tert* fibroblasts would be positively selected with puromycin. Puromycin titration was done as described in (Materials and Methods section 2.2.8.1). Briefly, 5 X $10^4$ cells/well were plated out in a 6-well plate cells and 24 hrs later puromycin was added. Cells were maintained in culture for 4 days post-plating before staining with crystal violet. The result (Figure 4.1) revealed that a concentration of 3 µg/ml puromycin was sufficient to kill BALB/c 3T3 *scid* fibroblasts.

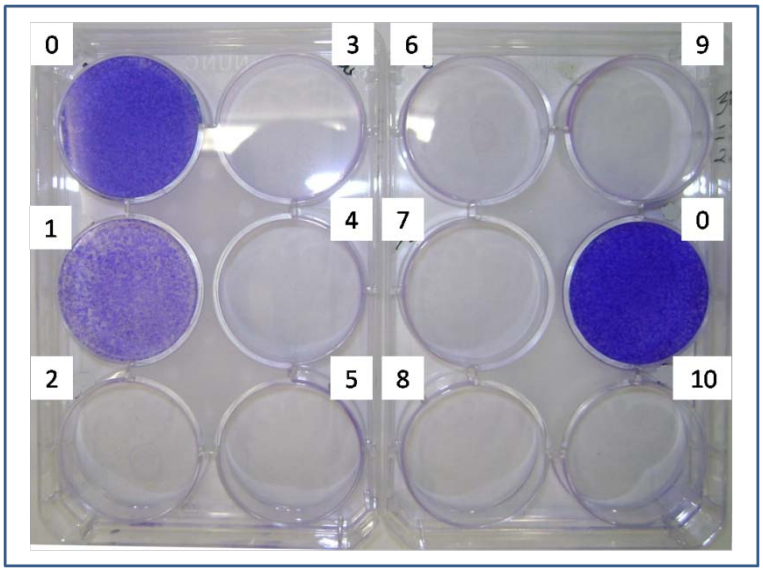



**Figure 4.1: Titration of puromycin killing in BALB/c 3T3 *scid* fibroblasts.** BALB/c 3T3 *scid* fibroblasts were treated from day 1 post-plating with concentrations of puromycin ranging 1-10 µg/ml (numbers shown in the figure denote puromycin concentrations). Cells were kept in culture until the untreated control reached confluency (4 days post-plating) and stained with crystal violet. 3µg/ml puromycin was adopted for selection.

### 4.1.2 Preparation and titration of retroviral vectors

A bicistronic retroviral vector encoding both m*Tert* and *Pac* was prepared as described in (Materials and Methods section 2.2.8.3). Retroviral vectors prepared for immortalisation were titrated in BALB/c 3T3 *scid* fibroblasts in presence of 8 µg/ml polybrene as mentioned earlier (Materials and Methods section 2.2.8.4). The effect of polybrene on these cells will be examined below in section 4.2. As shown in Figure 4.2, the results demonstrated that CFUs were too confluent to account in dilutions ($10^0$ to $10^{-3}$ dilutions) indicating high retroviral titre. For all other dilutions the following formula was used to calculate the retroviral titre: Retroviral (TU)/ml = number of CFUs x $10^5$ (cells/well at the time of vector addition) x 1/vector dilution. The calculations revealed that titre levels were 6 X $10^{10}$ and 5 X $10^{10}$ TU/ml on harvest day 1 and day 2 titres; respectively.

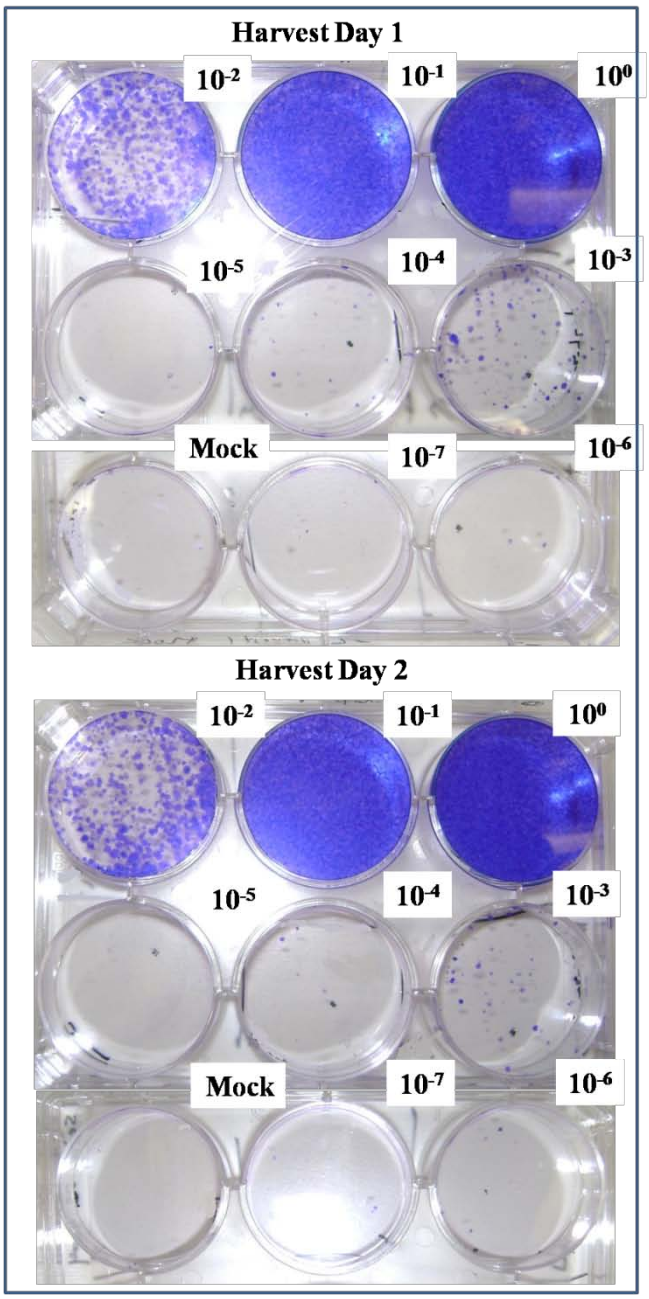



**Figure 4.2: Titration of retroviral vectors in BALB/c 3T3 *scid* fibroblasts**. Serial ten-fold dilutions ($10^{0}$ to $10^{-7}$) of retroviral vector supernatant of harvests day 1 or day 2 were prepared in 8 µg/ml polybrene. Two days post-transduction, 3 µg/ml of puromycin was added to each well and after three days the wells were stained by crystal violet to facilitate counting. The CFUs were TNTC in the dilutions of $10^{0}$ to $10^{-3}$ of both harvest days 1 and 2. For other dilutions, retroviral vector titres were calculated and the results showed that harvest day 1 and day 2 titres were 6 X $10^{10}$ and 5 X $10^{10}$ TU/ml; respectively.

### 4.1.3 Immortalisation by retroviral vectors

The prepared retroviral vectors were used to transduce BALB/c 3T3 *scid* fibroblasts in the presence of 8 µg/ml polybrene. The next day, cells were selected by adding (3 µg/ml) puromycin, and puromycin selection was maintained routinely for passaging of these cells. As a result of transduction, BALB/c 3T3 *scid* fibroblasts should be rendered puromycin resistant. In order to confirm that the transduced cells had acquired the resistance gene, the same puromycin titration was carried out as described above (section 4.1.1). The results in Figure 4.3 demonstrated that the fibroblast population was able to survive and resist to at least 10 µg/ml of puromycin, indicating that the BALB/c 3T3 *scid* fibroblasts has been transduced with the vector coding for *Pac* and m*Tert.*

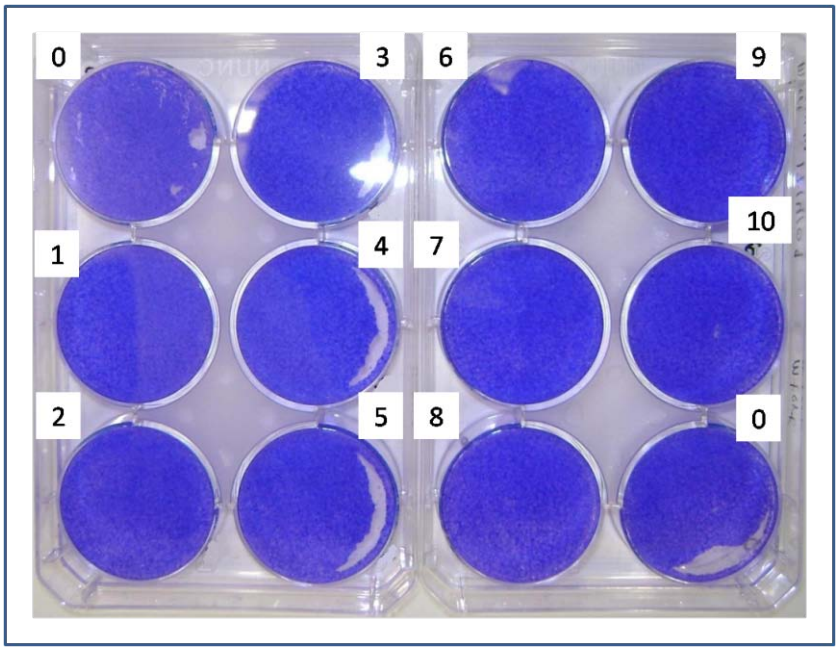



**Figure 4.3: Puromycin killing titration of BALB/c 3T3 *scid* m*Tert* fibroblasts.**

BALB/c 3T3 *scid* m*Tert* fibroblasts were grown in concentrations of puromycin ranged (1-10) µg/ml (numbers shown in the figure denote puromycin concentrations). Emerged puromycin-resistant cells were allowed to grow for 3 days post-plating to form a monolayer. Then, the wells were stained with crystal violet. The cells were able to grow in at least 10µg/ml puromycin.

### 4.1.4 Enhanced clonal efficiency and growth rate of m*Tert scid* fibroblasts

The ability of individual cells to form colonies in culture is called clonal efficiency, and can be measured by counting visible CFUs. Previous work (in Dr. Yáñez's laboratory) with BALB/c 3T3 *scid* fibroblasts had been impeded by cellular senescence. These cells showed low clonal efficiency and death before completion of the required experiments. In order to compare the clonal efficiency of immortalised and non-immortalised BALB/c 3T3 *scid* fibroblasts, an experiment was carried out by seeding low cell numbers of each and then assessment of CFUs (section 2.2.8.6). The results (see Figure 4.4) showed that for non-immortalised cells (BALB/c 3T3 *scid* fibroblasts) the CFU average was 4 and 35 for 10 and 100 seeded cells, respectively. While for immortalised cells (BALB/c 3T3 *scid* m*Tert* fibroblasts) the CFU average was 25 and TNTC for 10 and 100 seeded cells, respectively. This finding indicated that immortalised cells (*scid* m*Tert* fibroblasts) have significantly higher clonal efficiency than non-immortalised cells and they were more able to grow and expand i.e. possess higher clonal efficiency.

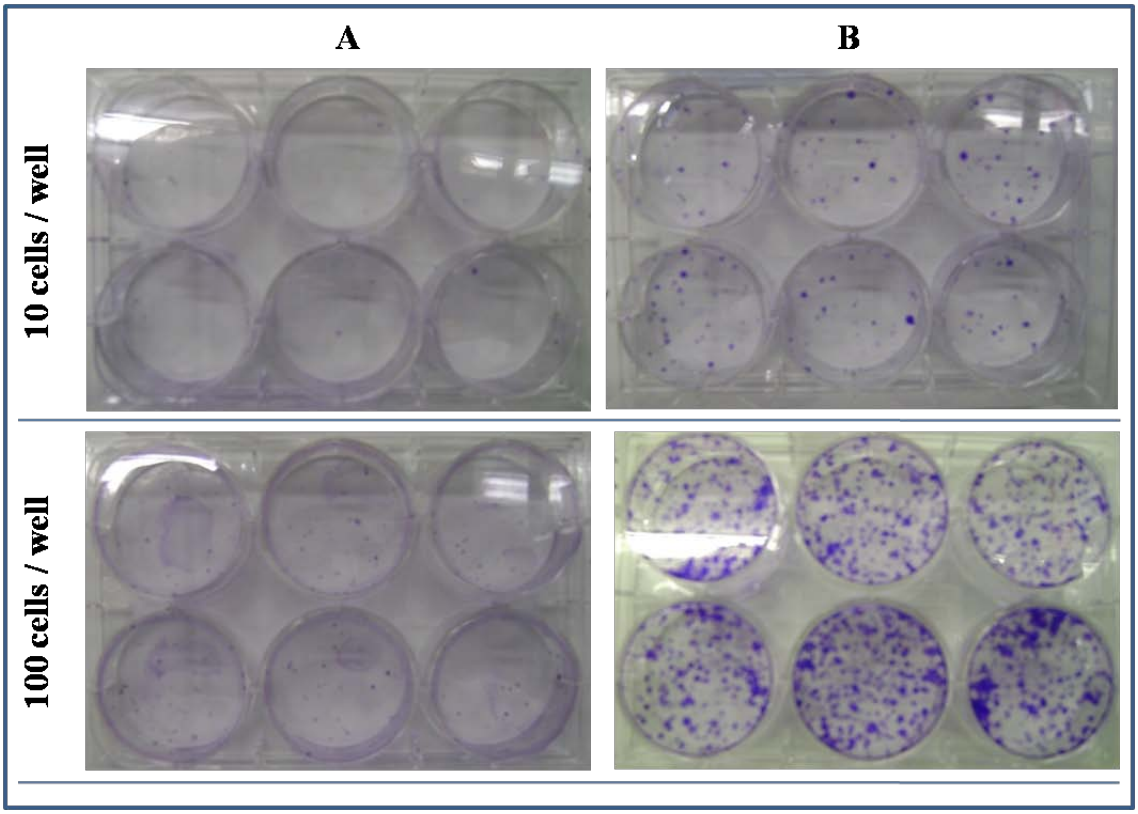



**Figure 4.4: Clonal efficiency of BALB/c 3T3 *scid* and *scid* m*Tert* fibroblasts.** Either 10 or 100 cells were seeded per well in 6-well plates. Four days later, emerged colonies were stained with crystal violet and then CFUs were scored. The results showed that CFU average was 4 and 35 for 10 and 100 seeded *scid* fibroblasts; or 25 and TNTC for *scid* m*Tert* fibroblasts, respectively. Clearly the *scid* m*Tert* fibroblasts have shown higher clonal efficiency than *scid* fibroblasts. Numbers on the left refers to the initial seeded cells/well. (A) BALB/c 3T3 *scid* fibroblasts; (B) BALB/c 3T3 *scid* m*Tert* fibroblasts (in the presence of 3 µg/ml puromycin).

To compare the growth kinetics of BALB/c 3T3 *scid* fibroblasts and *scid* m*Tert* fibroblasts, their growth curves were assessed (described in Materials and Methods section 2.2.8.6) by seeding specific cell numbers of each and then counting them in a haemocytometer at 24-hour intervals. The results showed that BALB/c 3T3 *scid* m*Tert* fibroblasts had divided noticeably faster (see Figure 4.5). In order to demonstrate significance, this experiment needs to be reapted.

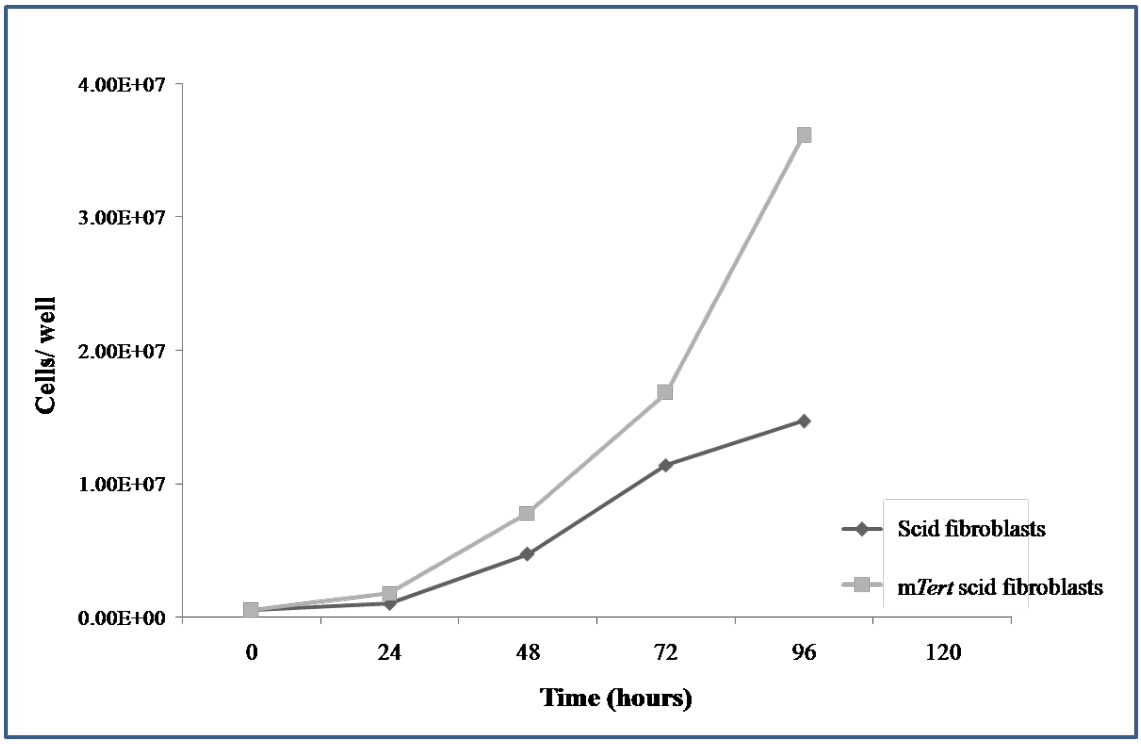



**Figure 4.5: Growth curve of mouse *scid* fibroblasts.** 5 X $10^5$ BALB/c 3T3 *scid* fibroblasts or *scid* m*Tert* fibroblasts were seeded per well in 6-well plates. The cells were maintained in culture and counted (by haemocytometer) at 24 hours intervals. BALB/c 3T3 *scid* m*Tert* fibroblasts divided faster than BALB/c 3T3 *scid* fibroblasts.

## 4.2 Polybrene effect on BALB/c 3T3 *scid* m*Tert* fibroblasts

Polybrene improves the transduction efficiency in some cells like HeLa and HEK-293T (Lehmusvaara S. *et al*., 2005) but not others like motor-neurons (Azzouz *et al*., 2004). In order to find out how BALB/c 3T3 *scid* m*Tert* fibroblasts will be affected by polybrene, an experiment was carried out using different eGFP MOIs of eGFP IPLVs and IDLVs. BALB/c 3T3 *scid* m*Tert* fibroblasts were transduced with lentiviral vectors either in presence or absence of 8 µg/ml of polybrene. The results of flow cytometry analysis (Figure 4.6) showed that the use of polybrene enhanced transduction efficiency for both IPLV and IDLV, with no obvious detrimental effect on the cells; therefore all transduction experiments in these cells have been done in the presence of polybrene. In order to demonstrate significance, this experiment needs to be reapted

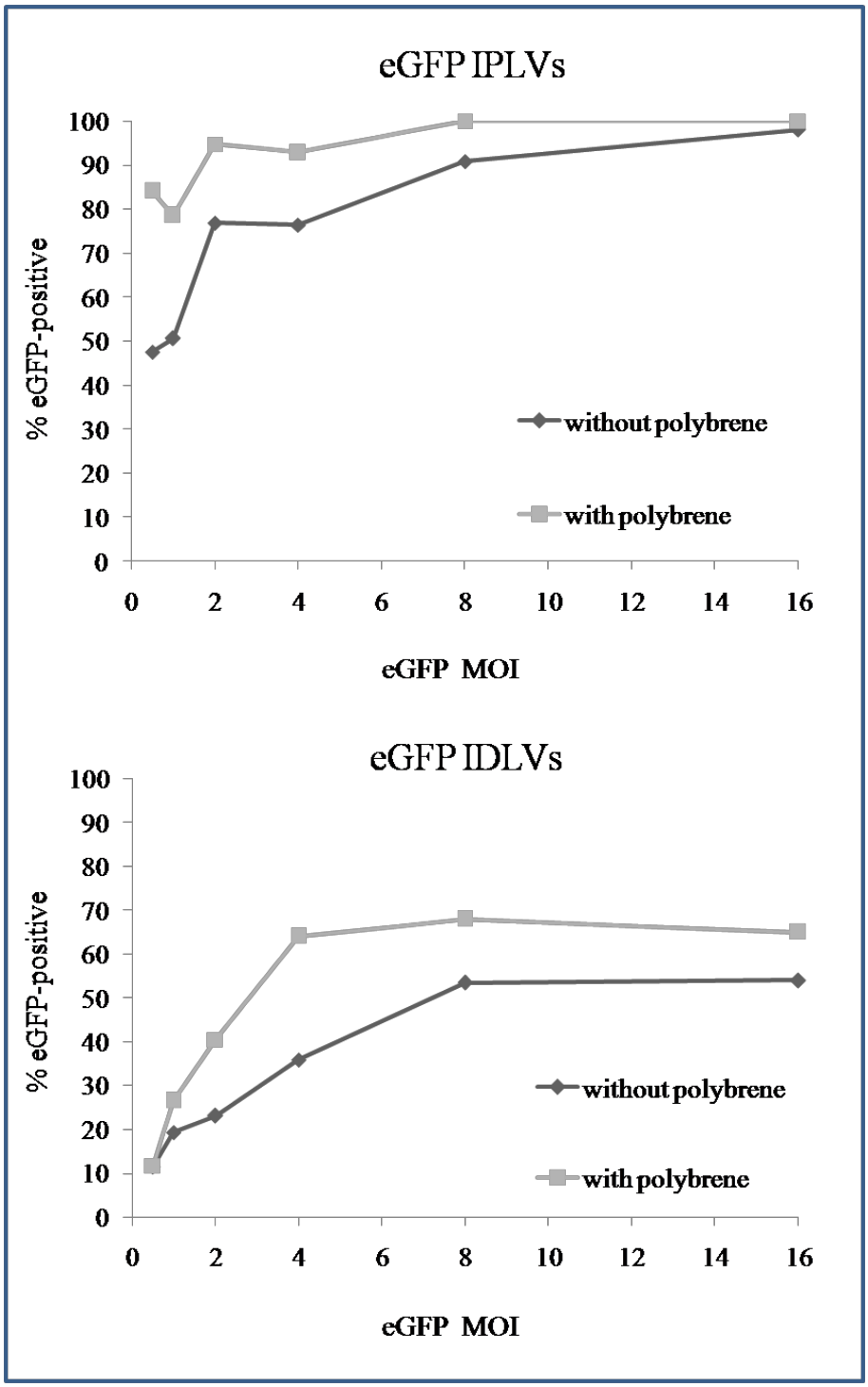



**Figure 4.6: Polybrene effect on transduction's efficiency of BALB/c 3T3 *scid* m*Tert* fibroblasts.** Cells were transduced at various MOIs with eGFP IPLVs and IDLVs in presence or absence of 8 µg/ml of polybrene. The transduction efficiency was estimated after three days, scoring eGFP-positive cells by flow cytometric analysis.

## 4.3 Optimisation of plasmid transfection

In current study, calcium phosphate coprecipitation (described earlier in Materials and Methods in section 2.2.5) was used for plasmid transfection, hence an optimisation of cell density and DNA amount was deemed required. Initially cell numbers of BALB/c 3T3 *scid* m*Tert* fibroblasts were tested; different cell numbers were seeded in 10 cm plates: ($5X10^5$, $1x10^6$, and $2x10^6$)**.** Then a fixed DNA amount of 10 µg/plate of eGFP-expressing plasmid (pRRLsc-CEW) was used for transfection. Three days later cells were analysed by flow-cytometry. The results (Figure 4.7) showed that all samples displayed similar percentages of eGFP-positive cells (8.3-12.4%). The $1x10^6$ cells per plate was chosen as it is more convenient cell density.

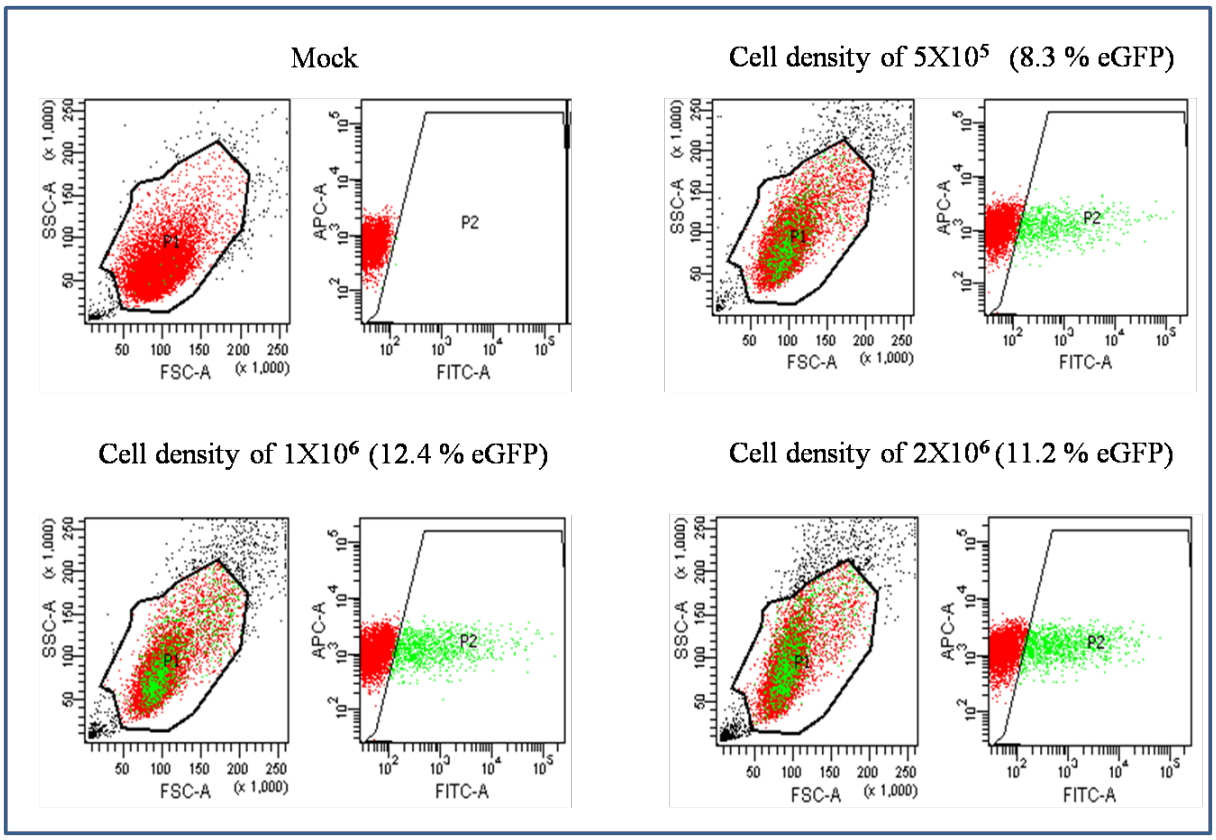



**Figure 4.7: Flow cytometric results of BALB/c 3T3 *scid* m*Tert* fibroblasts density optimisation for plasmid transfection.** Different cell densities were used ($5X10^5$, $1x10^6$, and $2x10^6$ cells per 10-cm plate), samples were then transfected with plasmid pRRLsc-CEW (10 µg/plate). Flow cytometric analysis showed that cell density of $1x10^6$ gave higher eGFP expression percentage of 12.4.

Furthermore, the ideal DNA concentration to be used for transfection was established using the selected density of $1x10^6$ cells/plate. Different amounts of pRRLsc-CEW were used: 0, 10, 20, 25, and 30 µg/plate. The flow cytometric results indicated that the percentages of cells expressing eGFP were 0, 9.4, 27.2, 28.5 and 28.3, respectively. Thus, a DNA amount of 20 µg/plate was considered optimal, as further increases did not lead to higher increase in eGFP expression.

## 4.4 Gene targeting in BALB/c 3T3 *scid* m*Tert* fibroblasts

### 4.4.1 Optimisation of amount of wt template plasmid

This experiment was conducted in order to optimise the required amount of wt homologous template plasmid p*Prkdc*Hind*Pml*I*neo*F in plasmid-based gene targeting experiments (below in section 4.4.2). Adjusting template plasmid amount was important to avoid cytotoxic or overdosing effect in the targeted cells. This gene targeting experiment used cell density of $1x10^6$ cell /plate of BALB/c 3T3 *scid* m*Tert* fibroblasts and 20 µg/plate of pRRLsc-CEW that were optimised in previous experiment (section 4.3). Additionally, to each plate an amount of 0, 5, 10, 15, 20, 25, or 30 µg/plate of p*Prkdc*Hind*Pml*I*neo*F were introduced by calcium phosphate coprecipitation followed by flow cytometric analysis after three days. In Table 4.1, the results showed that the DNA amount of 10 µg/plate of p*Prkdc*Hind*Pml*I*neo*F was the appropriate amount as adding more DNA does not lead to major increase in % eGFP.

**Table 4.1: Flow cytometric results of optimisation of p*Prkdc*Hind*Pml*I*neo*F amount.** All plates received equal amounts of 20 µg/plate of pRRLsc-CEW in addition to increasing amounts (as indicated in the table) of plasmid pPrkdcHindPmlIneoF by calcium phosphate coprecipitation. The flow cytometric results showed that the DNA amount of 10 µg/plate considered the most appropriate one.

| Sample (µg/plate of p*Prkdc*Hind*Pml*I*neo*F ) | % eGFP |
|---|---|
| Mock (non-transfected) | 0.0 |
| 0 | 27 |
| 5 | 29 |
| 10 | 39 |
| 15 | 40 |
| 20 | 33 |
| 25 | 33 |
| 30 | 30 |

### 4.4.2 Selection of optimum ZFN: Gene targeting using ZFN plasmids

This gene targeting experiment was based on transfection of BALB/c 3T3 *scid* m*Tert* fibroblasts with wild-type homologous corrective *Prkdc-neo* template plasmid (p*Prkdc*Hind*Pml*I*neo*F) and ZFN constructs. It was designed to detect which ZFN version (and combination) induces higher DSBs that can be detected by number of G-418 resistant CFUs. As I mentioned before, three versions of *Prkdc* ZFNs were designed for current study by Sangamo BioScience (California, USA) as a collaboration: the first contained wild-type *Fok*I domains which allow homodimerisation, the second contained modified wild-type *Fok*I domains which could allow some homodimerisation, and the third contained only obligatory heterodimeric *Fok*I domains.

The plasmid p*Prkdc*Hind*Pml*I*neo*F contains *Prkdc-neo* template of 9210 bp, with a *neo* cassette inserted at *Pml*I site in intron 84. *Neo* gene will confer resistance to G418 in gene-targeted cells or cells in which plasmid is randomly integrated (for more description see Figure 4.8). Where HR occurs, the region of homology of the corrective *Prkdc-neo* template will be used as a template for re-building endogenous genome sequence meaning that the *neo* cassette sequence will be incorporated into the genome. Upon positive selection with G418 and irrespective of the occurrence of site-specific gene repair or random integration, only those cells containing an integrated *neo* cassette would be able to generate visible G-418 resistant CFUs, which can be counted at the end of experiment.

G-418 is an analog of the antibiotic neomycin sulphate, its resistance is conferred by acquisition and expression the *neo* gene encoding the bacterial neomycin transferase (aminoglycoside 3' phosphotransferase, APH3'III) (Fernex *et al*., 1997). Neomycin transferase catalyzes the ATP-dependent phosphorylation of kanamycin/neomycin by transferring of the phosphoryl group of ATP to the 3'-hydroxyl of the antibiotic resulting in inhibition the antibiotic-ribosome interaction by blocking polypeptide synthesis at the elongation step in both prokaryotic and eukaryotic cells (Mondorf *et al*., 2012).

This experiment was performed as described in (section 2.2.10) in triplicate and repeated three times by plasmid co-transfection of both *Prkdc-neo* template plasmid p*Prkdc*Hind*Pml*I*neo*F and two ZFN plasmids of different versions and combinations. Table 4.2 summarises used ZFN versions, combination and added DNA amounts. For each plate the added DNA included: about 20 µg of ZFN plasmids (as optimised in section 4.3) or irrelevant

plasmid pCCLsc_ISceIT, and 10 µg/ plate of p*Prkdc*Hind*Pml*I*neo*F (as optimised in section 4.4.1). The irrelevant plasmid was used to make sure that all plates received approximately 32 µg of plasmid DNA. Latter plasmid was irrelevant to the ZFN and template plasmids and contained the same CMV promoter as that on ZFN constructs. The BALB/c 3T3 *scid* m*Tert* fibroblasts were seeded and allowed 4 hours for attachment, and then the cells were transfected with plasmid constructs. At two and five days post-transfection the culture medium was replaced with fresh medium containing 800 µg/ ml of G418. Then at 9 days post-transfection the plates were stained with crystal violet, and visible G-418 resistant CFUs counted.

The results (Figure 4.9) showed undoubtedly that the combination of obligate heterodimeric ZFN1 and ZFN2 (third version) generated the highest number of G-418 resistant (r) CFUs, as an indication of highest DSB generation. There was at least a 10-fold increase in G-418r CFUs between this combination and any other combination or negative control. Crucially, these findings were similar to what Olsen *et al* achieved earlier using very similar protocol. These authors had 9-fold increases in puromycin selected colonies when they transfected target cells with donor template and two separate ZFN monomers (Olsen *et al*., 2010).

Only those obligatory heterodimeric ZFN that generated the highest number of G-418r CFUs were incorporated into lentiviral vectors.

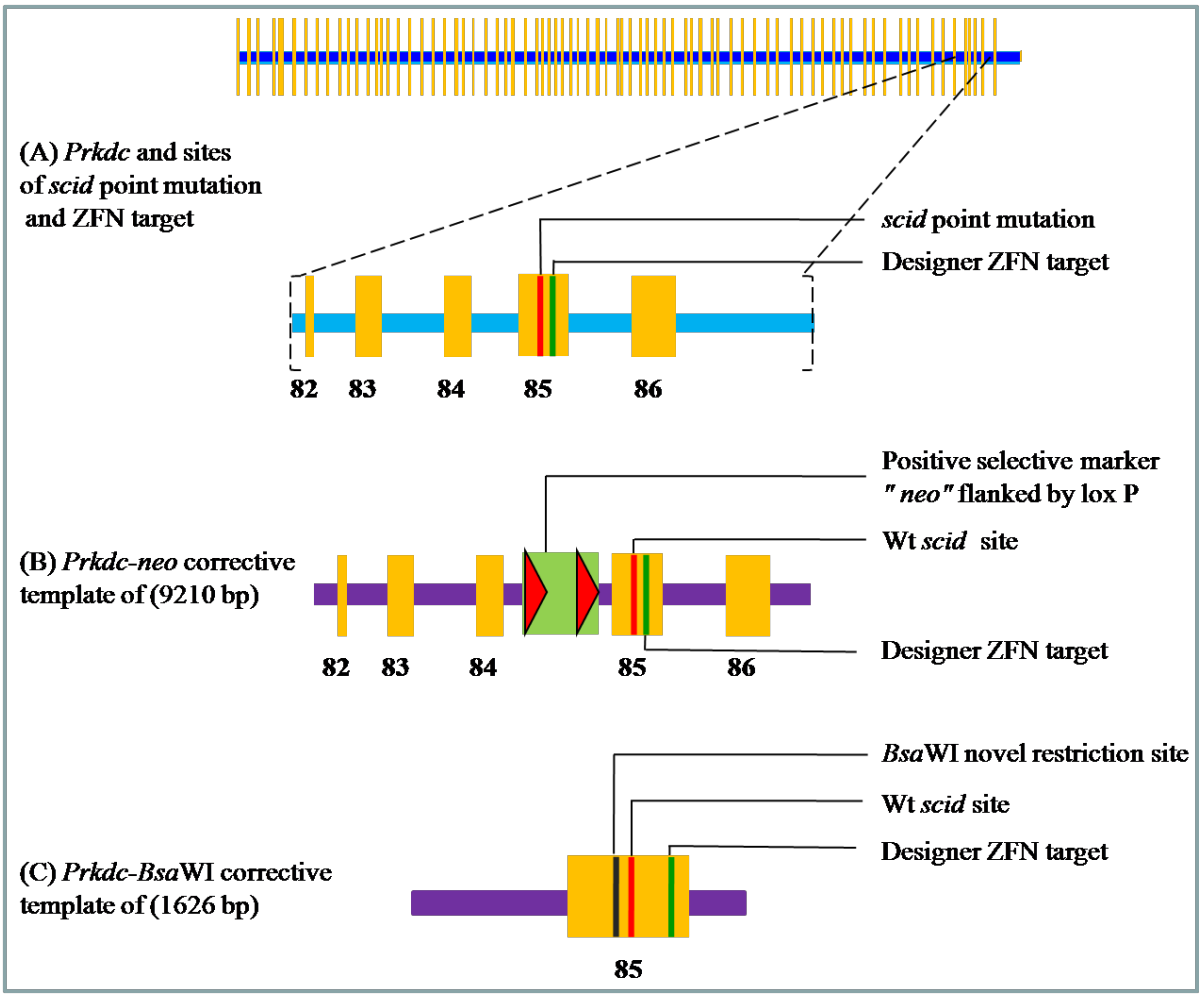



**Figure 4.8: General schematic of *Prkdc* gene and corrective templates.** (A) *Prkdc* configuration, comprising of 86 exons extending across 193 kb, and an enlarged *Hin*dIII fragment of *scid Prkdc* showing the location of the designer ZFN target site 32 nucleotides downstream from *scid* point mutation. (B) *Prkdc-neo* corrective template of 9210 bp. The floxed *neo* cassette was cloned into intron 84 for positive selection. Normal (wt) *scid* sequence and ZFN target sites are on the same 85 exon. (C) *Prkdc-Bsa*WI corrective template of 1626 bp containing a diagnostic *Bsa*WI restriction site introduced by site-directed mutagenesis at 2 nucleotides upstream from *scid* mutation site causing only silent changes, wt *scid* sequence and ZFN target sites 32 nucleotides downstream from *scid* mutation site.

**Table 4.2: Plan of gene targeting experiment for detection of efficient ZFNs.**

Three versions of *Prkdc* ZFNs were designed for current study; the first contained wild type non-optimised *FokI* domains which allow homodimerisation, the second was designed to include modified *FokI* domains which still allow some homodimerisation, and the third version contained only obligatory heterodimeric *FokI* domains. v: version.

| | | DNA added amounts (µg/plate) | | | |
|---|---|---|---|---|---|
| Plate # | Plate contents | ZFN plasmid (s) | Irrelevant plasmid (IP) pCCLsc_ISceIT | *Prkdc-neo* template plasmid pPrkdcHindPmlIneoF | Total added DNA (µg/plate) |
| 1 | Mock (non-transfected) | 0 | 0 | 0 | 0 |
| 2 | Negative control (without IP ) | 0 | 0 | 10 | 10 |
| 3 | Negative control (with IP ) | 0 | 22.2 | 10 | 32.2 |
| 4 | ZFN1 ($1^{st}$ v) | 11.4 | 10.8 | 10 | 32.2 |
| 5 | ZFN2 ($1^{st}$ v) | 10.8 | 11.4 | 10 | 32.2 |
| 6 | ZFN1+ ZFN2 ($1^{st}$ v) | 11.4 + 10.8 | 0 | 10 | 32.2 |
| 7 | ZFN1 ($2^{nd}$ v) | 11.1 | 11.1 | 10 | 32.2 |
| 8 | ZFN2 ($2^{nd}$ v) | 10.8 | 11.4 | 10 | 32.2 |
| 9 | ZFN1+ ZFN2 ($2^{nd}$ v) | 11.1 + 10.8 | 0.3 | 10 | 32.2 |
| 10 | ZFN1 ($3^{rd}$ v) | 11.1 | 11.1 | 10 | 32.2 |
| 11 | ZFN2 ($3^{rd}$ v) | 10.8 | 11.4 | 10 | 32.2 |
| 12 | ZFN1+ ZFN2 ($3^{rd}$ v) | 11.1 + 10.8 | 0.3 | 10 | 32.2 |

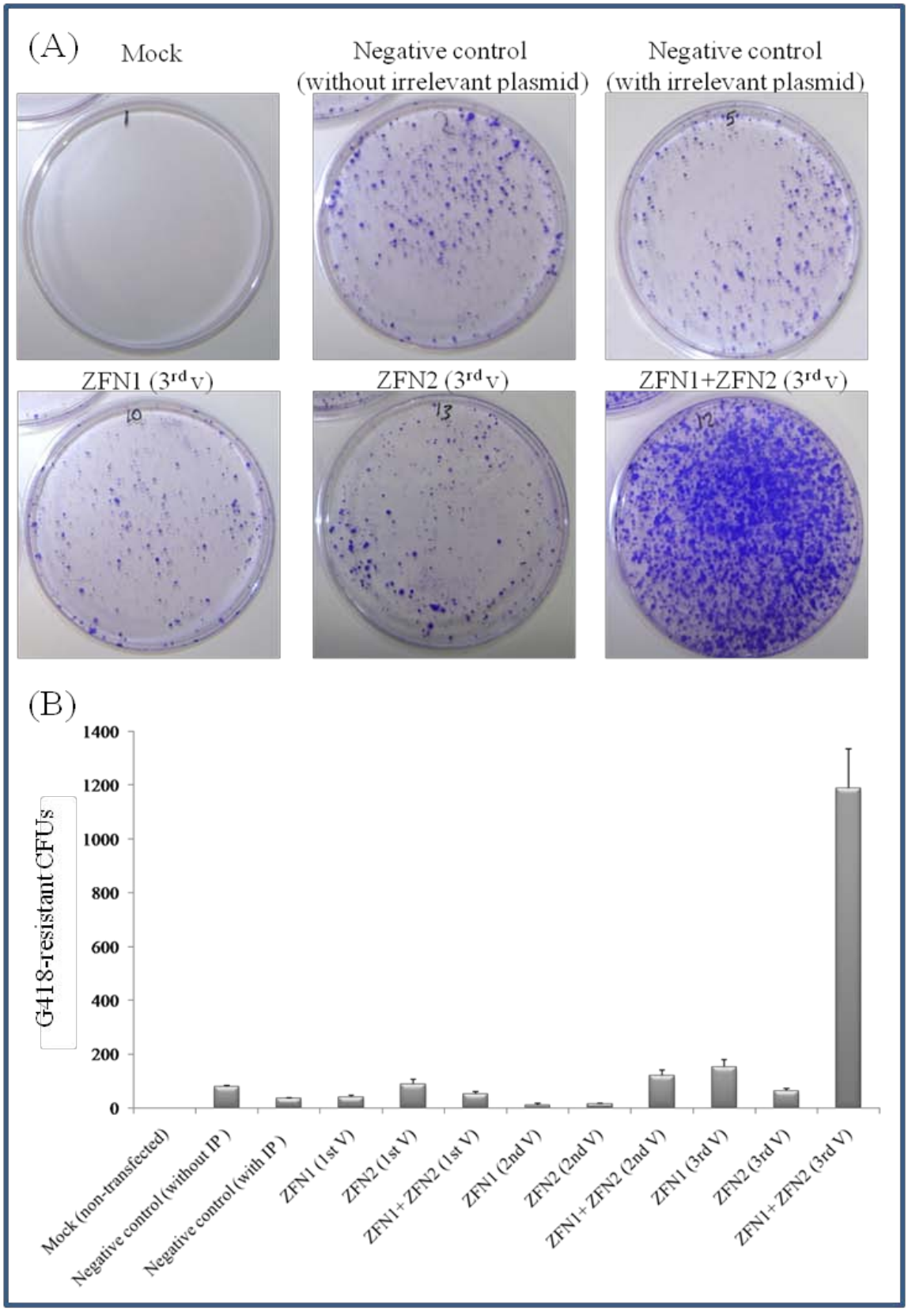

(A)
Mock
Negative control (without irrelevant plasmid)
Negative control (with irrelevant plasmid)
ZFN1 (3rd v)
ZFN2 (3rd v)
ZFN1+ZFN2 (3rd v)
(B)
G418-resistant CFUs
1400
1200
1000
800
600
400
200
0
Mock (non-transfected)
Negative control (without IP)
Negative control (with IP)
ZFN1 (1st V)
ZFN2 (1st V)
ZFN1+ZFN2 (1st V)
ZFN1 (2nd V)
ZFN2 (2nd V)
ZFN1+ZFN2 (2nd V)
ZFN1 (3rd V)
ZFN2 (3rd V)
ZFN1+ZFN2 (3rd V)

**Figure 4.9: Gene targeting in BALB/c 3T3 *scid* m*Tert* fibroblasts by plasmid transfection.** I utilised a plasmid-based gene targeting protocol using calcium phosphate coprecipitation of *Prkdc-neo* template plasmid and pVax-based ZFN plasmids. The BALB/c 3T3 *scid* m*Tert* fibroblasts were transfected with plasmid DNA (for quantities see Table 4.2). Two days later, the fibroblasts were washed with PBS and selected with G418 (800 µg/ ml). At 9 days post-transfection, the CFUs were stained with crystal violet and counted. Plates: Mock was non-transfected, negative control (no irrelevant plasmid) was transfected with only *Prkdc-neo* template plasmid (p*Prkdc*Hind*Pml*I*neo*F), negative control (with irrelevant plasmid) was transfected with p*Prkdc*Hind*Pml*I*neo*F and pCCLsc_ISceIT, ZFN1 and ZFN2 were the zinc-finger monomer plasmids, IP: irrelevant plasmid and v: version. (A) Representative photos of gene targeting experiment. There was a considerable increase in G-418 r CFU number when the ZFN obligate heterodimers were supplied together, in the presence of donor homologous template. (B) Quantitation of G-418r CFUs. The two obligate heterodimers together induced the highest number of G-418r CFUs, generating more than a 10-fold increase.

#### 4.4.2.1 Molecular analysis of G418r colonies

With the aim of demonstrating of ZFN specificity, a PCR analysis for targeted integration events was carried out on G418-resistant colonies from the previous experiment (section 4.4.2). 36 CFUs were picked randomly from two separate transfections of BALB/c 3T3 *scid* m*Tert* fibroblasts with corrective *Prkdc-neo* template (p*Prkdc*Hind*Pml*I*neo*F) and third version ZFN obligate heterodimeric (ZFN1 and ZFN2) plasmids.

The selected CFUs were allowed to grow into monolayers and then analysed by a PCR based protocol as described in section 2.2.10.1. The forward primer was designed to be internal to the targeting construct, and selected within the *neo* expression cassette sequences of the *Prkdc-neo* template plasmid (p*Prkdc*Hind*Pml*I*neo*F). The reverse primer was designed to be external to the *Prkdc-neo* template and selected from the genomic DNA immediately outside the short arm of homology (see Figure 4.10 for further description). PCR amplification of a 1335-bp product (for full sequence please see Appendix 1, Box 4) in tested samples is assumed to be diagnostic of ZFN-induced DSBs and HR (i.e. gene targeting) occurrence at these DSBs.

The results revealed that among 36 examined samples, there were 20 positive amplifications representing an impressive demonstration for ZFN-mediated gene targeting at the mutated *Prkdc* locus. The analysis of some representative sample clones is shown in Figure 4.11. Furthermore, to show that the right amplicons had been amplified these 20 positive products were subjected to enzymatic digestion with *Bgl*II, generating two smaller fragments of 746 and 589bp in all cases (see example in Figure 4.12 A). Additionally, among 20 positive amplicons, one sample was randomly selected to be subjected to further enzymatic digestion with *Hin*dIII, generating fragments of 1216 and 119 bp, and *Xho*I generating fragments of 1014 and 321 bp (see Figure 4.12 B).

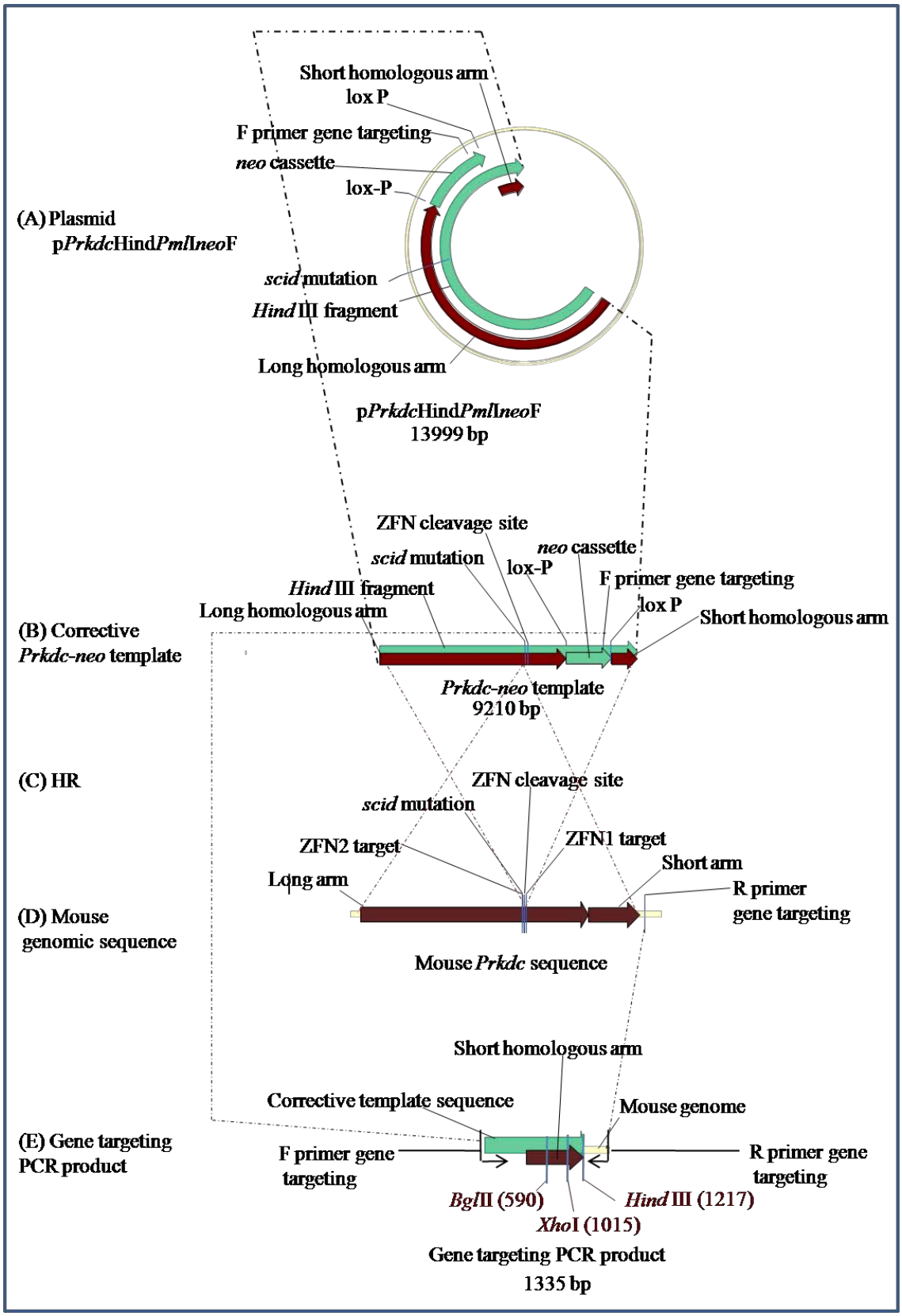

Short homologous arm
lox P
F primer gene targeting
neo cassette
lox-P
(A) Plasmid
pPrkdcHindPmlIneoF
scid mutation
Hind III fragment
Long homologous arm
pPrkdcHindPmlIneoF
13999 bp
ZFN cleavage site
scid mutation
neo cassette
lox-P
F primer gene targeting
Hind III fragment
Long homologous arm
lox P
Short homologous arm
(B) Corrective
Prkdc-neo template
Prkdc-neo template
9210 bp
(C) HR
ZFN cleavage site
scid mutation
ZFN2 target
ZFN1 target
Short arm
Long arm
R primer
gene targeting
(D) Mouse
genomic sequence
Mouse Prkdc sequence
Short homologous arm
Corrective template sequence
Mouse genome
(E) Gene targeting
PCR product
F primer gene
targeting
R primer gene
targeting
BglII (590)
Hind III (1217)
XhoI (1015)
Gene targeting PCR product
1335 bp

**Figure 4.10: Gene targeting strategy with *Prkdc-neo* targeting construct.** (A) Original plasmid p*Prkdc*Hind*Pml*I*neo*F containing the corrective *Prkdc-neo* template; (B) corrective *Prkdc-neo* template of 9210 bp encoding for wt (normal) sequence at *scid* mutation site, positive selection *neo* cassette and forward primer for gene targeting analysis; (C) proposed HR crossing-overs between corrective template and mouse genome; (D) *scid* mouse genome with *scid* mutation site and reverse gene targeting primer from outside the region of homology; and (E) PCR product of 1335 bp amplified using template–internal (forward) and template–external (reverse) primers, showing sites of *Bgl*II, *Xho*I and *Hind* III enzymes that will be used for subsequent confirmation.

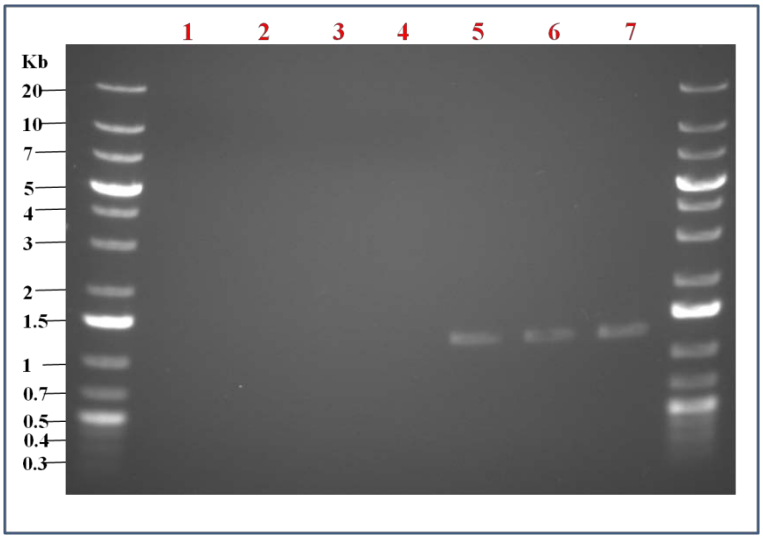



**Figure 4.11: Gene targeting at *Prkdc* locus in representative samples of BALB/c 3T3 *scid* m*Tert* fibroblast clones.** G-418 resistant CFUs resulting from potential gene targeting by ZFN and donor template plasmids were screened for HR by PCR amplification of a DNA sequence of 1335 bp extending from the *Prkdc-neo* corrective template into adjacent genomic DNA. Presence of positive amplification is diagnostic for gene targeting; out of 36 examined CFUs, 20 samples showed positive results. PCR products were run on a 0.7% agarose gel; lane 1 mock (non-transfected), lanes 2, 3 & 4 (clone numbers 3, 10 & 13) were samples failed to show amplification and lanes 5, 6 & 7 (clone numbers 1, 27 & 29) were samples with positive amplification. DNA ladder: 5 µl of Fermentas GenRuler 1 Kb plus.

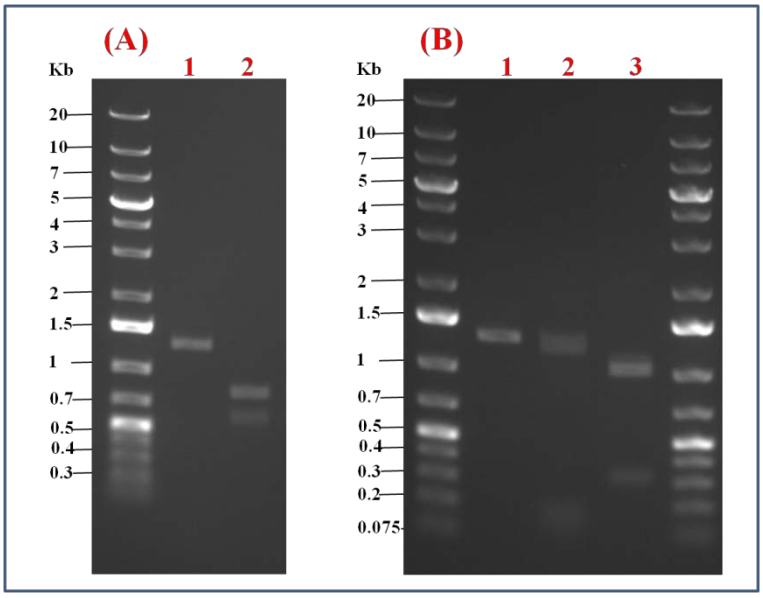



**Figure 4.12: Restriction endonuclease analysis of PCR amplicons of gene targeting.**

PCR products from gene targeted BALB/c 3T3 *scid* m*Tert* fibroblasts were digested with selected restriction enzymes to verify the presence of correct amplicons. Digested products were run on a 0.7% agarose gel. (A) All 20 PCR products were digested with *Bgl*II; one representative sample (clone number 1) is shown here. Lane 1 undigested product of 1335 bp; and lane 2 *Bgl*II-generated smaller fragments of 746 and 589 bp.
(B) One sample (clone number 5) was randomly selected out of 20 PCR products to be digested with another two restriction enzymes; lane 1 undigested product of 1335 bp, lane 2 digestion with *Hind*III generating fragments of 1216 and 119 bp, and lane 3 digestion with *Xho*I generating fragments of 1014 and 321 bp.

### 4.4.3 Gene targeting using ZFN IPLVs

Following the success of the all-plasmid gene targeting experiment described in the previous section, I set out to study whether gene targeting could be achieved through transduction with IPLVs encoding ZFN monomers driven by either CMV or SFFV; and transfection of *Prkdc-neo* template plasmid p*Prkdc*Hind*Pml*I*neo*F. The aim here was to verify the functionality of ZFNs expressed from lentiviral vectors. As I mentioned above (section 4.4.2), only obligate heterodimeric ZFNs were incorporated into lentiviral vectors, as they generated the highest number of G-418 resistant CFUs in targeted fibroblasts. Because I was unsuccessful in incorporating the *Prkdc-neo* template corrective plasmid into lentiviral vectors, ZFNs and *Prkdc-neo* template were introduced into the targeted fibroblasts by transduction and transfection, respectively. Different time intervals between ZFN transduction and template plasmid transfection were set up to give the ZFN lentiviral vectors varying times for gene expression ahead of plasmid template transfection.

This experiment was carried out in the presence of 8 µg/ml polybrene using the cell density of $1X10^6$ cells/plate. Four hours post-seeding, the fibroblasts were transduced with qPCR MOI 50 each of ZFN IPLV pairs (ZFN1 and ZFN2) driven by either CMV or SFFV promoters. Later the *Prkdc-neo* corrective template plasmid was introduced by plasmid transfection at different time intervals of 12, 24, 48 or 72 hours post-transduction. Two days after the corresponding transfection, the cells were selected with G418, and eventually the G-418 resistant CFUs were stained with crystal violet and counted. As shown in Figure 4.13, the results revealed clearly that ZFN IPLVs were able to express ZFNs efficiently producing G418 resistant CFUs as an indicator of induced DSBs. Additionally, transfection at 48 hours post-transduction was the most suitable time-point, judging by the number of G418-resistant colonies, and expression levels from CMV and SFFV ZFN IPLVs were similar in the targeted fibroblasts.

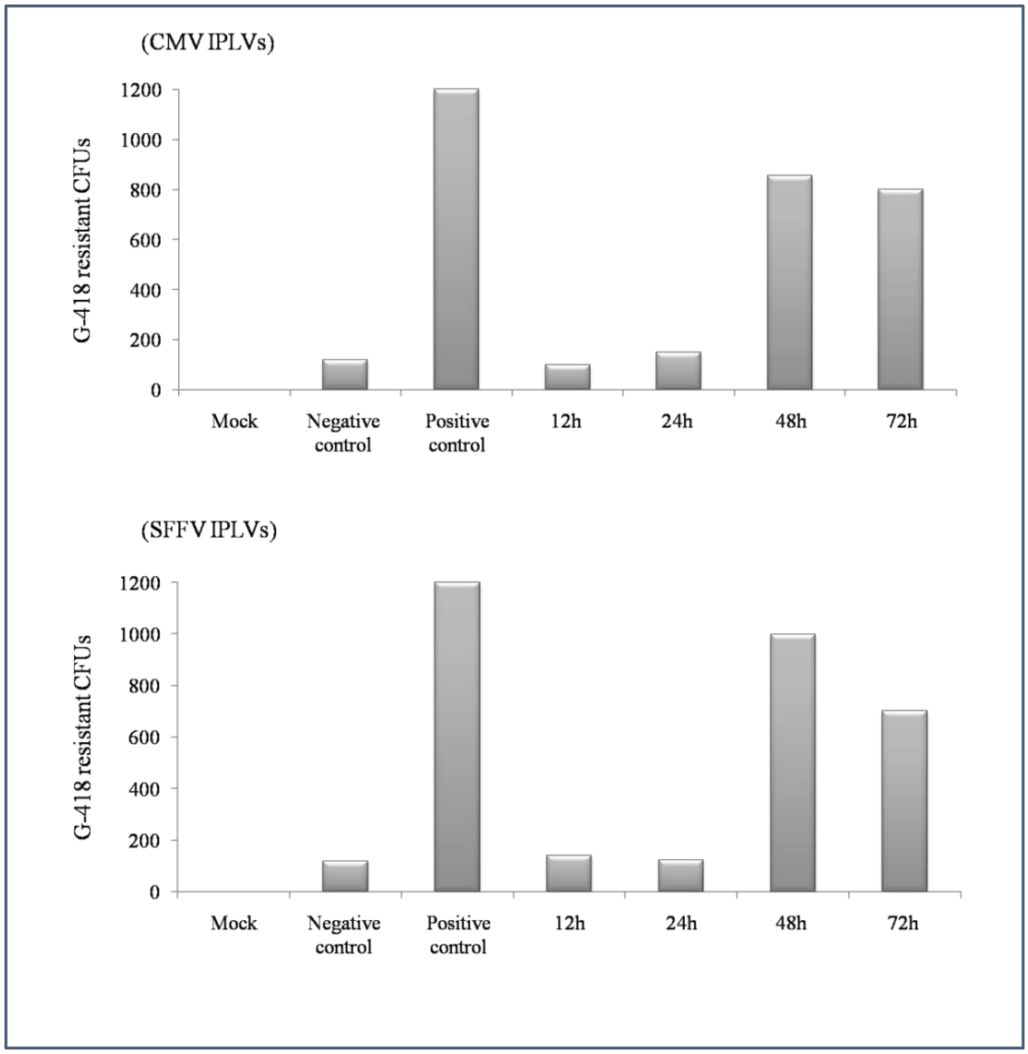



**Figure 4.13: Quantitation of G-418 resistant CFUs in BALB/c 3T3 *scid* m*Tert* fibroblasts following combined ZFN IPLV transduction and plasmid template transfection.** The fibroblasts were transduced with qPCR MOI 50 of ZFN IPLVs driven by either CMV or SFFV promoters in the presence of 8 µg/ml polybrene. Later, at 12, 24, 48 or 72 hours post-transduction, the fibroblasts were transfected with the corrective *Prkdc-neo* template plasmid. Two days after the corresponding transfection, the cells were selected with G418 and eventually the G-418 resistant CFUs were stained with crystal violet and counted. The ZFN IPLVs driven by CMV or SFFV were able to express ZFNs efficiently and an interval of 48 hours between transduction and transfection was the optimum timing. (Upper panel) Transduction with CMV ZFN IPLVs; (lower panel) Transduction with SFFV ZFN IPLVs. Negative control was corrective *Prkdc-neo* template plasmid only, while positive control was transfection with obligate heterodimeric (ZFN1 & ZFN2) and corrective template plasmids.

## 4.5 Testing the efficiency of ZFN cutting

In order to study ZFN abilities to generate site-specific cuttings at their genomic target site (32 nucleotides downstream from *scid* point mutation at the *Prkdc* locus), a Surveyor *Cel*-I mutation detection assay was used (section 2.2.9). In the absence of homologous template, ZFN-induced DSBs would be repaired by NHEJ, resulting in the introduction of InDels (small insertions or deletions) at high frequency. These InDels are detected by the Surveyor *Cel*-I nuclease assay (see Figure 4.14). The Surveyor nuclease detects mismatches in heteroduplex DNA produced in the assay, cleaving both strands of a DNA heteroduplex on the 3'-side of the mismatch site (Qiu *et al*., 2004). The Surveyor's *Cel*-I assay was used as described earlier (Materials and Methods section 2.2.9). Any polymorphic mismatches or introduced InDels (via repaired DSBs), eventually, will be recognised by *Cel*-I nuclease. Similarly, the detection will include those modifications introduced at the ZFN target site upon repair by NHEJ machinery in absence of any supplied donor homologues templates.

### 4.5.1 Sensitivity of *Cel*-I assay

This experiment was designed in order to find out the minimum modification percentage that can be detected by the Surveyor's *Cel*-I assay (described in Materials and Methods section 2.2.9.1). Having detected the minimum amount of ZFN cutting by this assay, will be crucial to optimise ZFN cleavage efficiency for the next experiments. As the only expected difference between BALB/c 3T3 wt fibroblast and BALB/c 3T3 *scid* fibroblast DNAs is the *scid* point mutation (T- to -A change), by mixing different amounts of each it should be possible to generate different proportions of homoduplex and heteroduplex DNA molecules. This was carried out by PCR amplification of a 304-bp region around the *scid* mutation from mixed genomic DNA (for full sequence please see Appendix 1, Box 1), followed by a final denaturation/hybridisation step. Any heteroduplexes would be detected and cut by the Surveyor's nuclease at the mispaired site (the *scid* mutation), generating two smaller fragments of 88 and 216 bp. The results of this test (see Figure 4.14) demonstrated that the minimum percentage of modification (mismatch) that could be detected by the *Cel*-I assay was 5%.

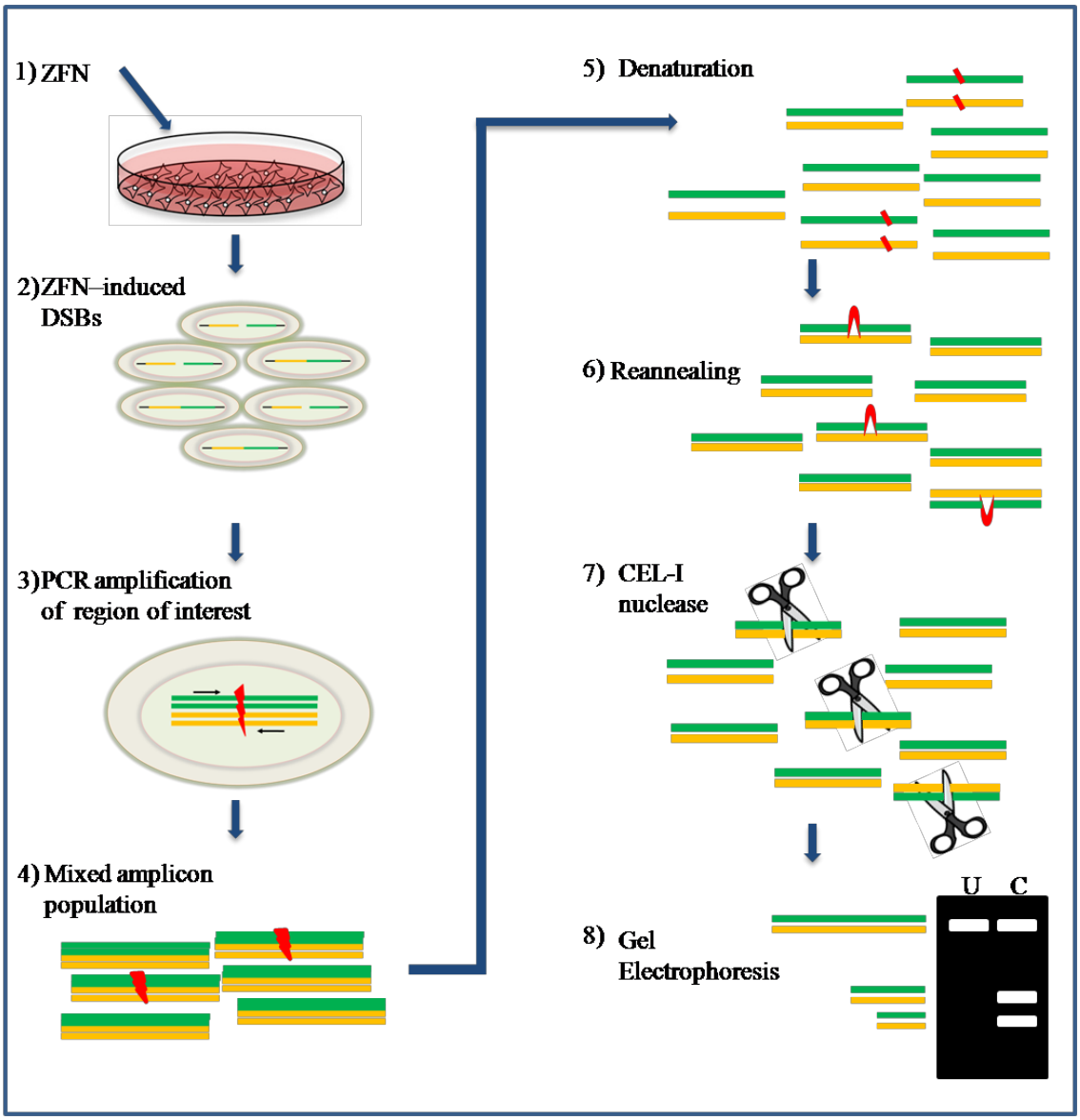



**Figure 4.14: Surveyor's *Cel*-I endonuclease detects ZFN activity**. (1) Transduction of cells with ZFN lentiviral vectors (2) ZFNs recognise their target sites creating DSBs in a portion of the cells. (3& 4) Those DSBs repaired by NHEJ, which will produce small insertions or deletions (InDels, ), DNA extraction from targeted cells and PCR amplification of target sites. (5) Denaturation of PCR products (6) Heteroduplex formation between normal and mutated products by reannealing. (7) Heteroduplex cleavage by Surveyor endonuclease. (8) Separation of Surveyor digests by electrophoresis and analysis of band sizes and proportions. U: Uncut, C: Cut.

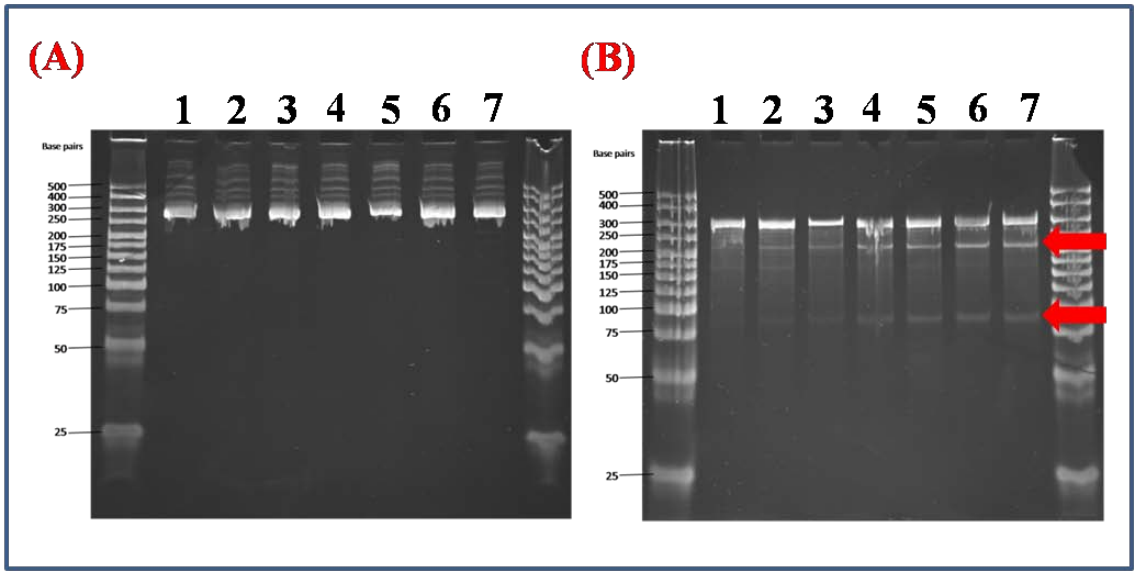



**Figure 4.15: Detection limit of Surveyor's *Cel*-I assay.** Different percentages of BALB/c 3T3 wt and *scid* m*Tert* fibroblasts were mixed, genomic DNA purified and DNA fragments of 304 bp around the *scid* mutation were PCR amplified. The PCR products were then subjected to the *Cel*-I assay, in which cutting by *Cel*-I nuclease (at *scid* point mutation) produced two smaller fragments of 88 and 216 bp (red arrows) in addition to the parental amplicon. The minimal percentage of modification that was detected by this assay was 5%. (A) PCR amplification and (B) Surveyor's *Cel*-I nuclease cutting. Lane (1) 99% of wt and 1% of *scid*, lane (2) 95% of wt and 5% of *scid*, lane (3) 90% of wt and 10% of *scid*, lane (4) 80% of wt and 20% of *scid*, lane (5) 70% of wt and 30% of *scid*, lane (6) 60% of wt and 40% of *scid* and lane (7) 50% of wt and 50% of *scid*. Marker: Bioline Hyperladder V.

### 4.5.2 Testing the efficiency of ZFN cutting in *scid* m*Tert* fibroblasts

The BALB/c 3T3 *scid* m*Tert* fibroblasts were transduced with either CMV or SFFV ZFN IPLVs and IDLVs at qPCR MOI 50, 125 or 250, and maintained in culture for three days. The cells were then collected, genomic DNA was extracted and subjected to the Surveyor *Cel*-I mutation detection assay. A PCR amplification across the target site generated a 304-bp parental fragment (for DNA sequence please see Appendix 1, Box 1). Denaturation followed by reannealing of this PCR product generated DNA heteroduplexes if ZFN-induced InDels were present. After cleavage by Surveyor's nuclease, two smaller products of 184 and 126 bp (Figure 4.16) were generated. Then samples were run in PAGE, visualized under UV light, images were taken and finally cutting percentages were calculated with ImageJ. The results showed that there was efficient dose- and dimer-dependent ZFN-mediated cutting at the target site, suggesting that ZFN expression under CMV or SFFV promoters reached similar levels. Furthermore and as expected in these proliferative cells, the expression level from IPLVs was more efficient than that from IDLVs for both CMV or SFFV-controlled ZFN cassettes.

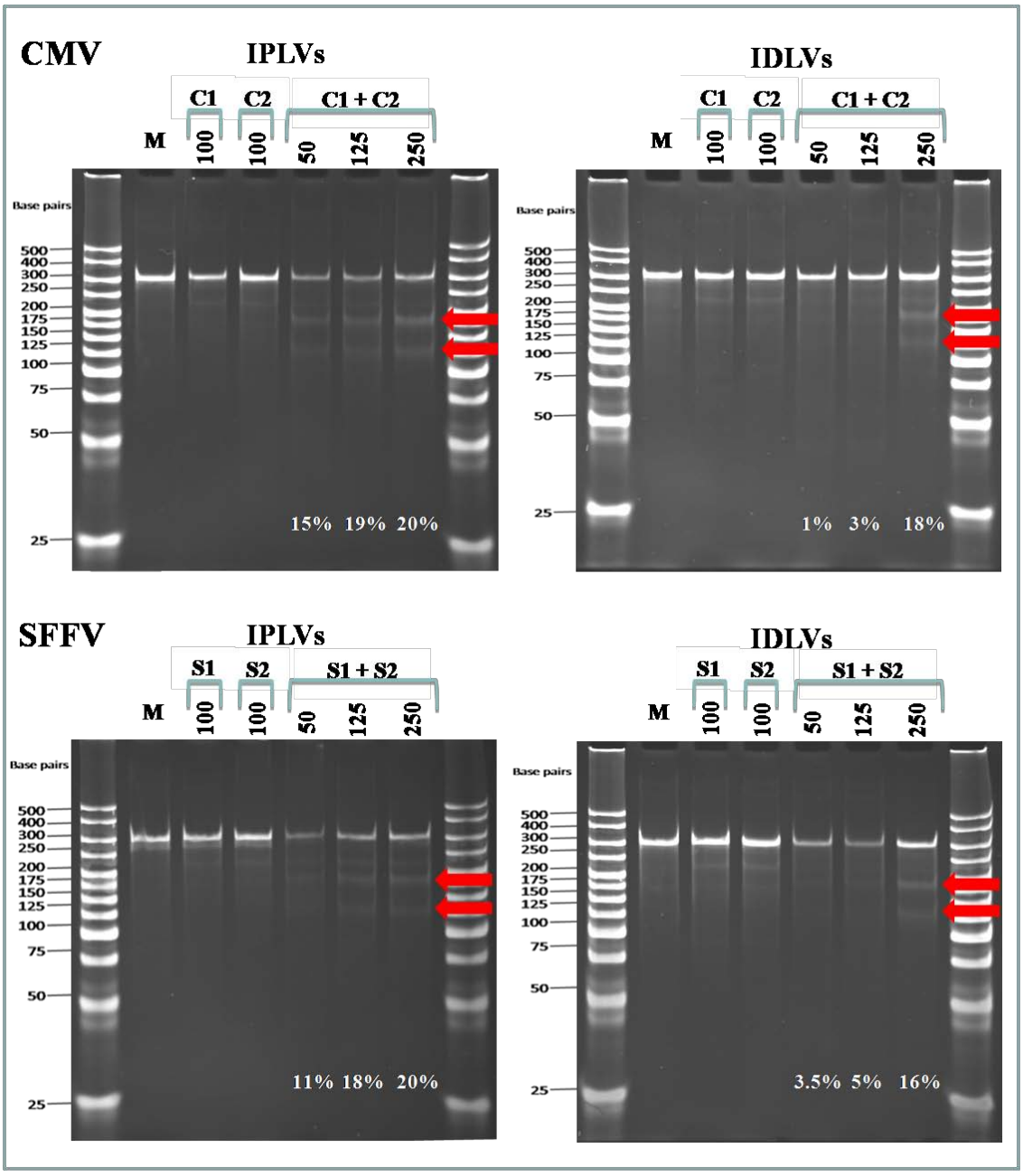

CMV
IPLVs
C1
C2
C1 + C2
M
100
100
50
125
250
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25
15% 19% 20%
IDLVs
C1
C2
C1 + C2
M
100
100
50
125
250
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25
1% 3% 18%
SFFV
IPLVs
S1
S2
S1 + S2
M
100
100
50
125
250
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25
11% 18% 20%
IDLVs
S1
S2
S1 + S2
M
100
100
50
125
250
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25
3.5% 5% 16%

**Figure 4.16: Site-specific ZFN cutting in *scid* fibroblasts.** BALB/c 3T3 *scid* m*Tert* fibroblasts were transduced at the indicated qPCR MOIs of 50, 125 or 250 with either CMV ZFN lentiviral vectors (upper panel) or SFFV ZFN lentiviral vectors (lower panel). Three days later cells were harvested, genomic DNA was extracted and then samples were subjected to Surveyor *Cel*-I mutation detection assay. The PCR amplification generates a 304-bp parental fragment, and after denaturation/reannealing and cleavage by Surveyor's nuclease, two smaller fragments of 184 and 126-bp (red arrows) are produced wherever InDels are present. The results demonstrated site-specific cutting, suggested similar levels of ZFN expression under CMV or SFFV promoters, and that the expression level from IPLVs was more efficient than that from IDLVs for both CMV or SFFV-driven IPLVs. Percentages at the bottom of each image represent ZFN cutting percentage calculated with ImageJ. M: mock, C1: CMV ZFN1, C2: CMV ZFN2, S1: SFFV ZFN1 and S2: SFFV ZFN2.

## 4.6 Detection of gene repair by *Bsa*WI assay in *scid* m*Tert* fibroblasts

The aim behind this experiment was to detect HR events in targeted fibroblasts using a silent diagnostic restriction site for the enzyme *Bsa*WI, previously introduced via site-directed mutagenesis two nucleotides upstream from the *scid* mutation site in the *Prkdc-Bsa*WI corrective template. The *scid* fibroblasts were transduced with lentiviral vectors encoding the *Prkdc-Bsa*WI corrective template of 1626 bp and either CMV or SFFV driven ZFNs. The sequence of the *Prkdc-Bsa*WI corrective template was verified by DNA sequencing (for full sequence please see Appendix 1, Box 2). Upon gene repair, the supplemented *Prkdc-Bsa*WI template including the diagnostic *Bsa*WI site will be used as a template to repair the genomic locus. The *Bsa*WI assay is based on PCR amplification across the target site using primers external to the targeting construct, amplicon's digestion with *Bsa*WI and then Southern blotting to detect the generated bands (see Figure 4.17 for more description). This work has been done in collaboration with Dr. Céline Rocca at RHUL.

In this context two Southern blot probes were prepared as described (section 2.2.12.5.2), both from the same *Prkdc-Bsa*WI template but differing in their length: a long probe of 1708 bp and a short probe of 1068 bp (see Figure 4.18). The first and longer was the full PCR amplicon of 1708 bp, amplified using the primers external to the *Prkdc-Bsa*WI template, purified and electrophoresed to check size and purity. On other hand, the shorter one was prepared by *Bsa*WI digestion of the long probe of 1708 bp, preparative electrophoresis, retrieval of the 1068-bp fragment from the gel and purification. Firstly, the 1708 bp probe was used, but because the smaller band (of 640 bp) detected on the gel tends to overlap with non-specific smear, making the calculation of gene repair percentage potentially inaccurate, the shorter probe of 1068 bp was used instead.

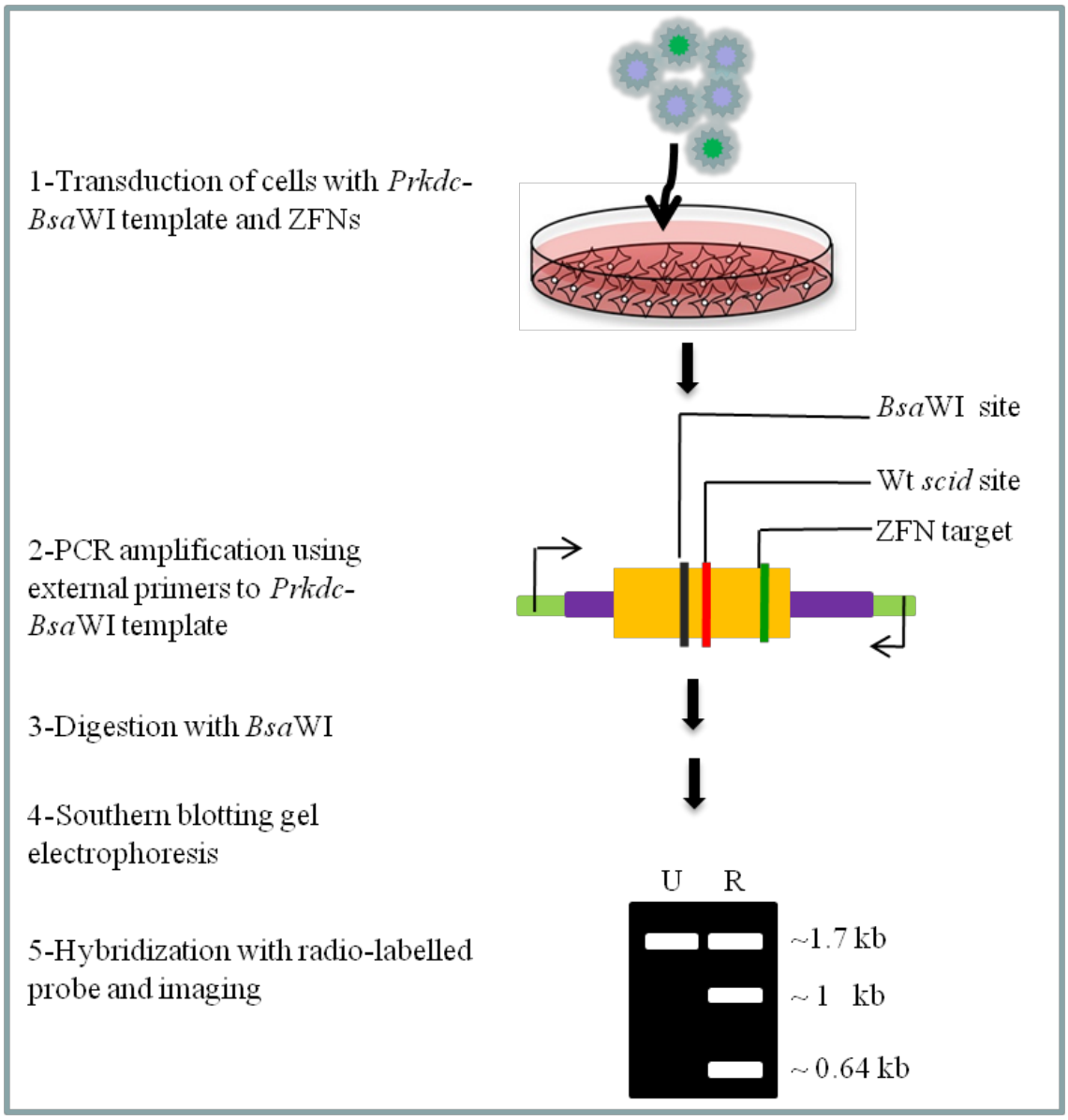



**Figure 4.17: *Bsa*WI assay detects ZFN-mediated gene repair**. (1) Transduction of the m*Tert scid* fibroblasts with *Prkdc-Bsa*WI template and ZFN lentiviral vectors (2) Genomic DNA extraction from targeted cells and PCR amplification of region of interest (1708 bp) using external primers to *Prkdc-Bsa*WI template (3) Digestion overnight (~ 16 h) with *Bsa*WI (4) Electrophoresis of the samples on Southern blotting gel and then transfer onto nylon membrane (5) Hybridization with radio-labelled probe and imaging. U: Unrepaired, R: Repaired.

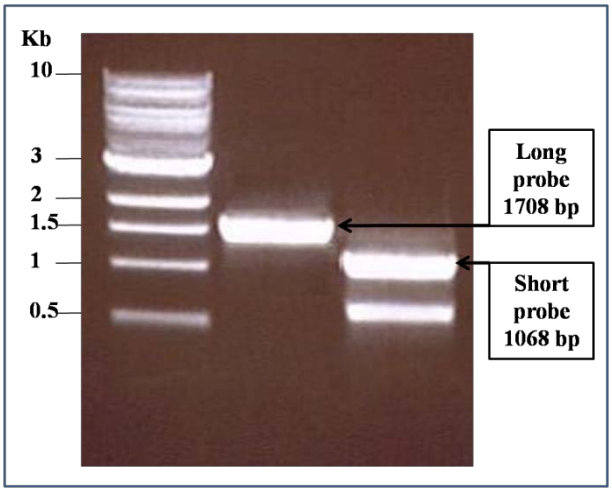



**Figure 4.18: Long and short probes for *Bsa*WI assay on 1% agarose gel.** The long probe (1708 bp) was amplified using primers external to the corrective *Prkdc-Bsa*WI template (1626 bp), purified, and electrophoresed to check for expected size and purity; whereas the shorter one (1068 bp) was prepared by *Bsa*WI digestion of the long probe, electrophoresed, retrieved from the gel and finally purified. DNA ladder: 5 µl of NEB 1 Kb ladder.

In order to estimate HR at the *Prkdc* locus in this *Bsa*WI assay, BALB/c 3T3 *scid* m*Tert* fibroblasts were transduced as described earlier (section 2.2.12.2). Many transduction scenarios were tested using different MOIs, ratios and combinations of ZFNs and *Prkdc* template. Remarkable results were observed with two transduction plans. The first was a three-vector transduction used for ZFN IPLVs (ZFN1 and ZFN2) and *Prkdc-Bsa*WI template IDLV at a ratio of ZFN1: ZFN2: template (1: 1: 2). While the second was a two-vector transduction used for *Prkdc*/ZFN IDLVs, where each vector encoded the *Prkdc-Bsa*WI template and one of the ZFN monomers, keeping the molar ratio of ZFN1: ZFN2: template (1: 1: 2). Tests were conducted in which ZFNs were driven by either CMV or SFFV on both IPLV and IDLV configurations.

Ten days post-transduction, cells were split into two samples: the first was kept aside for further analysis (see next sections) and the second used here for this *Bsa*WI assay. For this part, the genomic DNA was extracted and the *scid* target region was PCR amplified using external primers to the corrective *Prkdc-Bsa*WI template. Next, the amplicons were digested overnight with *Bsa*WI. The digestion products were

electrophoresed, transferred onto nylon membrane, hybridised with radio-labelled probe, and the blot was scanned and imaged. Finally, gene repair percentages were determined using ImageJ by quantification of each band's intensity (I), 1708 and 1068-bp bands were imaged, and the following formulas were applied:

When the long (1708-bp) probe was used, molecules of the full-length 1708-bp uncut band are more intense than those of the 1068-bp band, and a correction factor is needed. The correction factor enhances the intensity of the 1068-bp band to make it equal to that of the 1708-bp band. Correction factor = 1708/1068=1.60.

Gene repair % = $(100 \text{ X } 1.6 \text{ X } I_{1068}) / (I_{1708} + [1.6 \text{ X } I_{1068}])$.

While If the small (1068-bp) probe was used, hybridising molecules of the full-length 1708-bp uncut band or the 1068-bp band produce the same intensity. No correction factor is needed in this case.

Gene repair % = $(100 \text{ X } I_{1068bp}) / (I_{1708bp} + I_{1068bp})$.

A non transduced (mock) sample was used as a negative control, while a positive control was prepared by PCR amplification using the plasmid pHind-TK*Bsa*WI (containing a fragment of *Prkdc* around the *scid* site and the *Bsa*WI target sequence introduced by site-directed mutagenesis) as a template. The positive control was amplified using the same primers that were used previously to amplify long probe of 1708 bp (Material and method section 2.2.12.3). Additional negative controls were prepared by transduction of the cells with IDLVs encoding either template only (MOI 10000) or one ZFN monomer only (MOI 5000). The results revealed (see Figure 4.19) that gene repair can be detected when SFFV driven ZFNs were used in both IPLVs and IDLVs, as unequivocal detection of the 1068 bp was diagnostic for gene repair. Furthermore, gene repair was vector dose-dependent. Surprisingly, no diagnostic band was noticed in the case of CMV driven ZFNs used in both IPLVs (data not shown) and IDLVs.

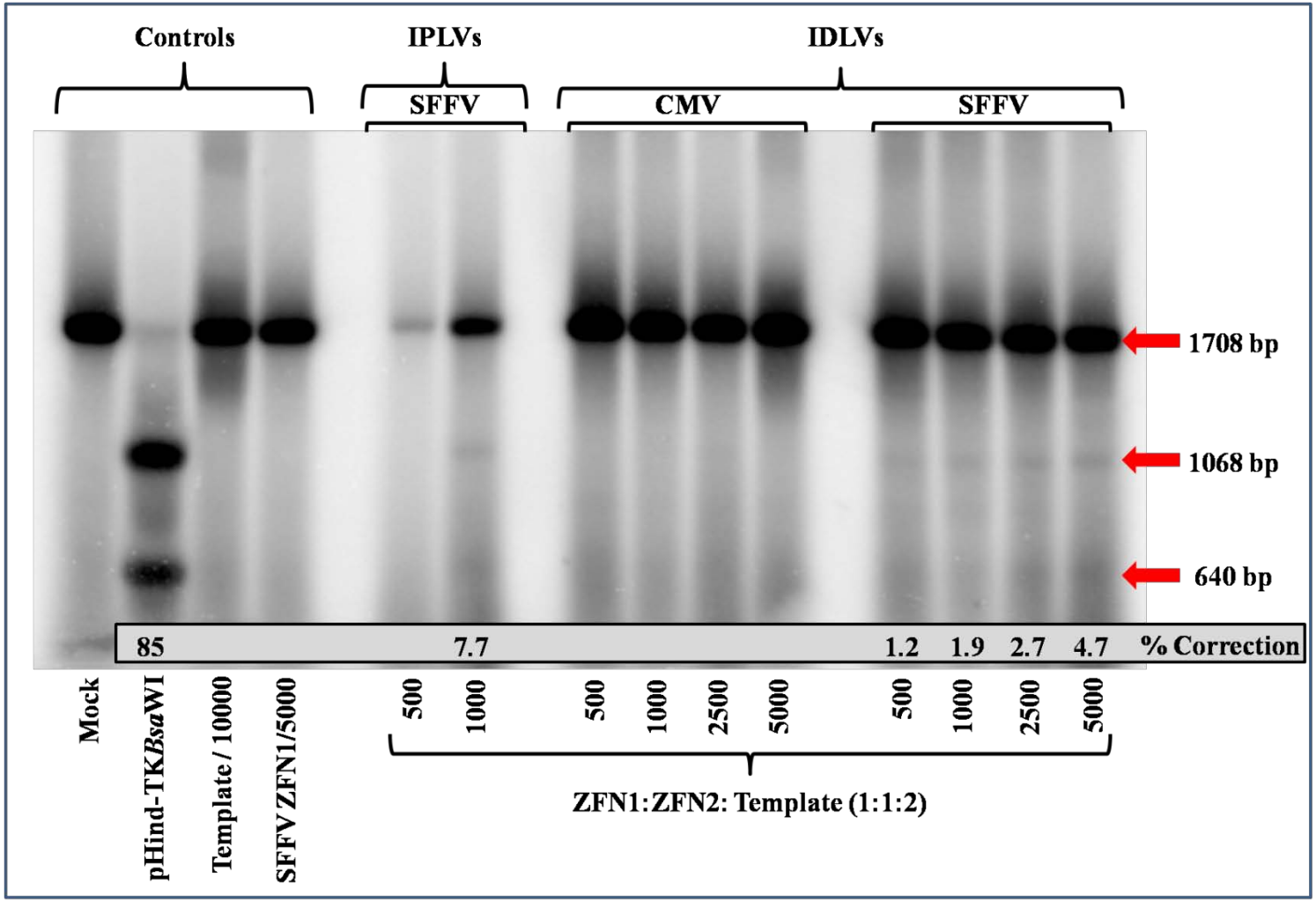

Controls
IPLVs
IDLVs
SFFV
CMV
SFFV
1708 bp
1068 bp
640 bp
85
7.7
1.2
1.9
2.7
4.7
% Correction
Mock
pHind-TK*Bsa*WI
Template / 10000
SFFV ZFN1/5000
500
1000
500
1000
2500
5000
500
1000
2500
5000
ZFN1:ZFN2: Template (1:1:2)

**Figure 4.19: Gene repair at *Prkdc* locus in BALB/c 3T3 *scid* m*Tert* fibroblasts measured by *Bsa*WI assay.** Ten days post-transduction, the *Bsa*WI assay was used to detect gene repair events. The presence of the 1068-bp band was diagnostic for gene repair. Gene repair percentages were determined using ImageJ by quantification of band intensities as described in main text using the long (1708-bp) probe. Gene repair was detected when SFFV-driven ZFNs were used in both IPLVs and IDLVs; whereas no correction was noticed in case of IDLVs ZFNs driven by CMV. **Controls**: mock was non-transduced fibroblasts, pHind-TK*Bsa*WI was PCR amplification of original plasmid containing the template modified by site directed mutagenesis, template/10000 and SFFV ZFN1/5000 were transduced fibroblasts with IDLVs encoding for *Prkdc-Bsa*WI template (MOI 10000) and SFFV ZFN1 (MOI 5000); respectively. **IPLV samples**: fibroblasts were transduced at the indicated MOIs with IPLVs driven by SFFV. Three vectors were used: ZFN1 and ZFN2 IPLVs; and *Prkdc-Bsa*WI template IDLVs at ratio of ZFN1: ZFN2: template (1: 1: 2). For example the number 500 at the bottom means that the cells were transduced with MOI 500 of each ZFN IPLV and MOI 1000 of *Prkdc-Bsa*WI template. **IDLV samples**: fibroblasts were transduced at the indicated MOIs with IDLVs driven by either CMV or SFFV. Two vectors were used: template/ZFN1 and template/ ZFN2 IDLVs at molar ratio of ZFN1: ZFN2: template (1: 1: 2). For example the number 2500 at the bottom means that the cells were transduced with MOI 2500 of each of template/ZFN1 and template/ ZFN2 IDLVs.

## 4.7 Rescue of DNA-PK activity

Gene repair of the mutated genotype should lead to restoration of functional phenotype; in this context correction of *scid* point mutation in the *Prkdc* gene is expected to restore the DNA-PK activity as *Prkdc* encodes this kinase. Therefore; this experiment was designed to investigate whether the activity of DNA-PK was restored after gene repair of *Prkdc* was observed in section 4.6. In order to do that DNA-PK activity was measured using SignaTECT DNA-dependent protein kinase assay as described earlier (section 2.2.13). This work has been done in collaboration with Dr. Céline Rocca.

In brief, those samples that showed positive gene repair results in the above *Bsa*WI assay (i.e. SFFV ZFN IPLVs and IDLVs) were collected at ten days post-transduction. Non-transduced BALB/c 3T3 *scid* m*Tert* fibroblasts were used as a negative control (mock), whereas wt BALB/c 3T3 fibroblasts were used as a positive control. Nuclear proteins from the selected samples were extracted using CelLytic nuclear extraction kit. Later, these nuclear proteins were incubated with a biotinylated p53-derived DNA-PK peptide substrate, an activation buffer containing double strand DNA and [$\gamma$-$^{32}$P] ATP. Next, the biotinylated peptide substrate was captured by spotting reactions on individual squares on the high binding capacity capture membrane characterized by a high density of streptavidin. After several washes of the membrane, the DNA-PK activity was quantified by two different systems for more accuracy: scintillation counter and phospho-imaging system. Finally the specific activity of DNA-PK was calculated as explained earlier (section 2.2.13.2) in pmol ATP/minute and µg of protein. This assay was repeated twice.

In Figure 4.20, the results are plotted as a ratio of enzyme specific activity of DNA-PK in tested samples relative to that in mock (negative control). There was a vector dose-dependent restoration of DNA-PK activity when *scid* fibroblasts were transduced with corrective *Prkdc-Bsa*WI template IDLVs and SFFV driven ZFN IPLVs or IDLVs. The enzymatic activity of gene-repaired samples was obviously lower than that of wt BALB/c 3T3 fibroblasts, due to the fact that the tested samples were polyclonal populations (of mutated and repaired cells) in which only a small proportion was gene-repaired.

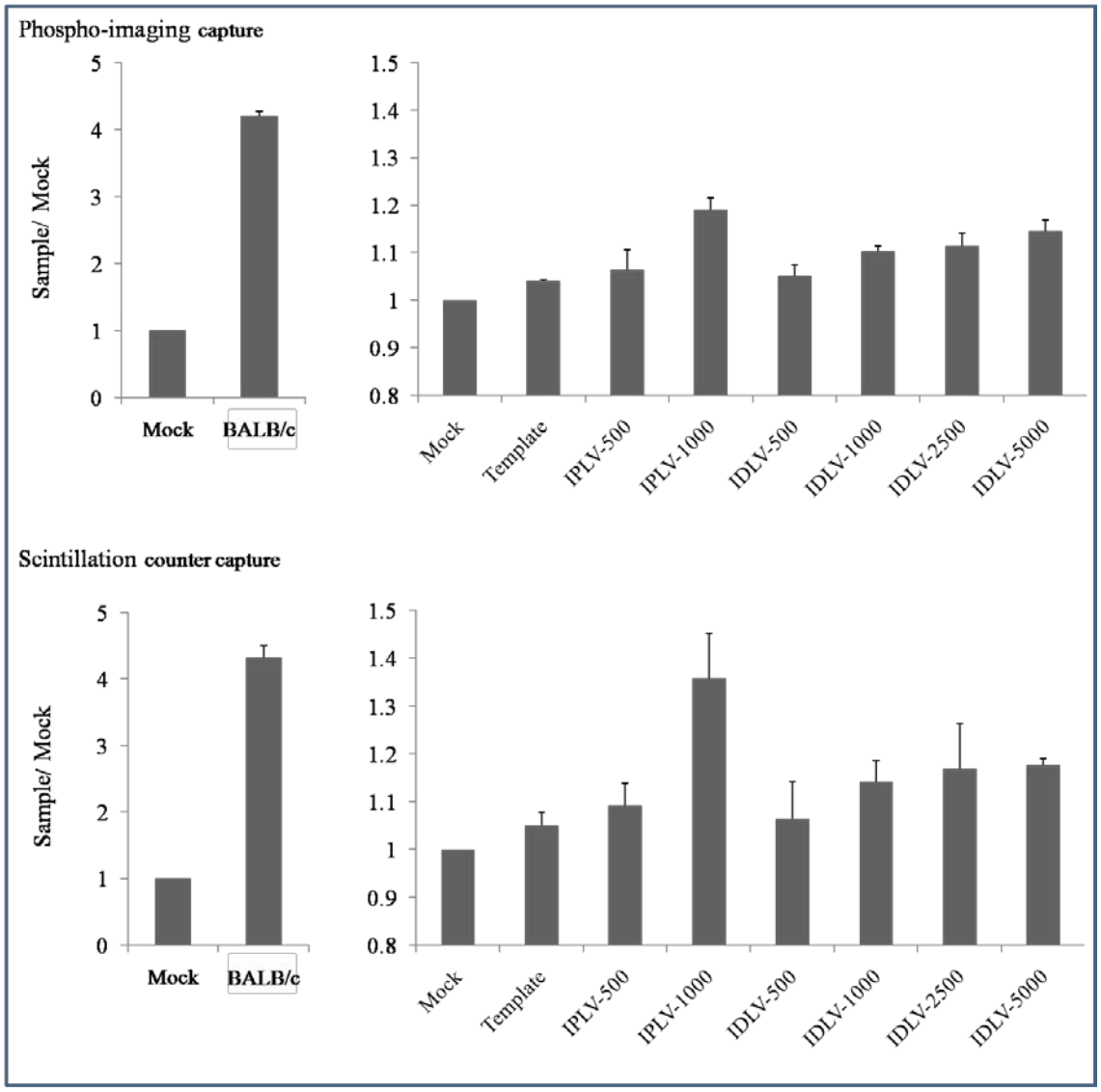



**Figure 4.20: DNA-PK specific activity in gene-repaired BALB/c 3T3 *scid* m*Tert* fibroblasts.** *Scid* fibroblasts were transduced with corrective *Prkdc-Bsa*WI template IDLVs and SFFV-driven ZFN IPLVs or IDLVs at the indicated MOIs. Ten days post-transduction, nuclear proteins were extracted by CelLytic nuclear extraction kit and DNA-PK activity of samples was measured by SignaTECT DNA-dependent protein kinase assay. The results showed that there was a vector dose-dependent restoration of DNA-PK activity which correlated with gene repair.

## 4.8 Enrichment of corrected cells

Rescue of *Prkdc* functionality in corrected cells should restore the ability to respond to DNA damaging agents, this being further evidence of phenotype rescue of *scid* cells after gene repair. Gene-repaired polyclonal populations from the *Bsa*WI assay (section 4.6) were subjected melphalan treatment to assess their phenotype correction. This part of the work was carried out in collaboration with Dr. Céline Rocca.

### 4.8.1 Determination of optimum melphalan dose

To optimise the cytotoxic dose of melphalan on wt and *scid* fibroblasts, the MTT assay was used to measure cell viability at different melphalan concentrations. The optimum dose for enrichment of corrected cells would be chosen to maximise killing of DNA-PK deficient cells (mutated *scid* cells) and survival of DNA-PK proficient cells (gene repaired cells).

As described earlier in section 2.2.14.1, the BALB/c 3T3 wt and *scid* m*Tert* fibroblasts were separately seeded and in the next day different melphalan concentrations ranged (0.01 -50 µM) were added for 1 hour. After 5 days, MTT was added and left for 4 hours, and then DMSO was added and only 100 µl from each sample were used to measure the optical density at 570 nm. The viability of the cells was worked out using that optical density at melphalan zero concentration, which was considered as 100% viability. The results showed differential viability of wt and *scid* fibroblasts, as expected. The main finding in this experiment was that a concentration of 10 µM of melphalan provided the maximum differential, with approximately 80 % of DNA-PK deficient cells and 40 % of DNA-PK proficient cells being eliminated (see Figure 4.21). These findings suggested that 10 µM melphalan could be used to attempt to enrich polyclonal populations of gene-repaired cells.

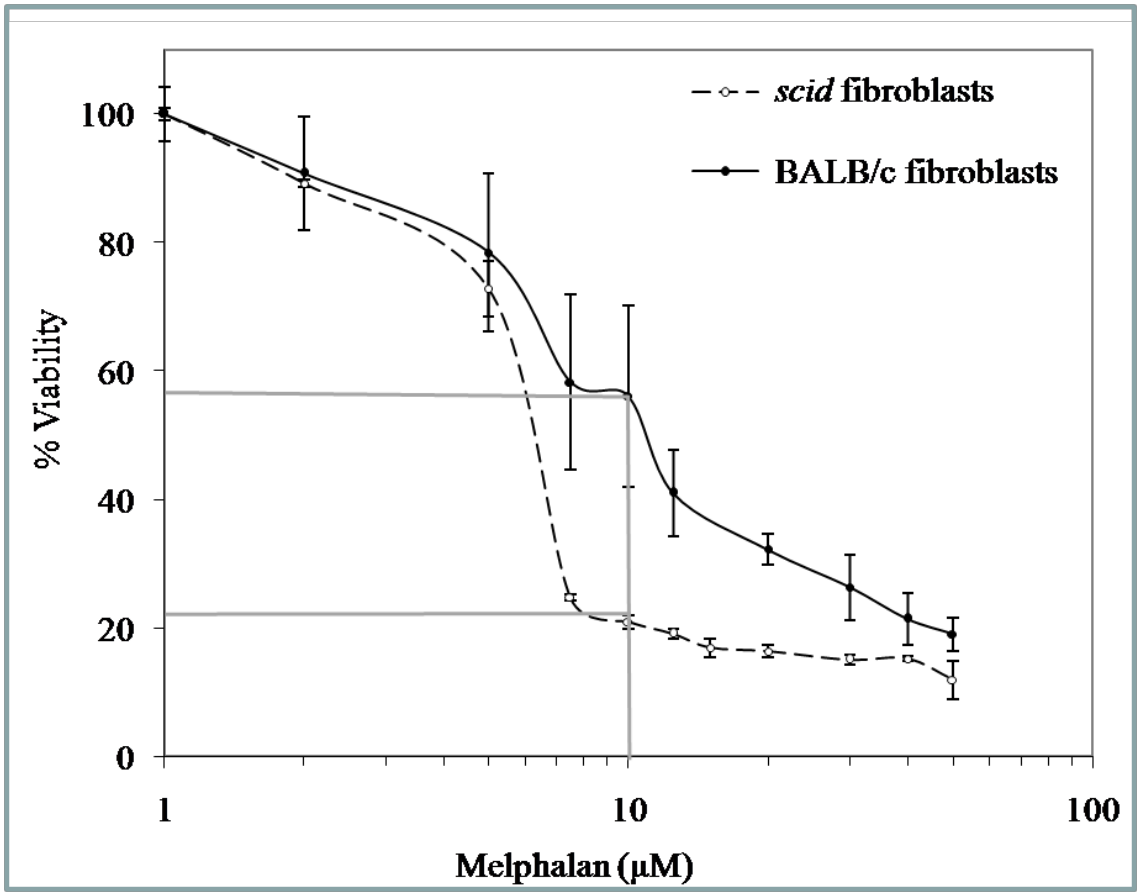



**Figure 4.21: Determination of optimum cytotoxic dose of melphalan in BALB/c 3T3 wt and *scid* fibroblasts measured by MTT assay.** The fibroblasts were seeded, and the next day the culture medium was replaced with fresh medium containing different melphalan concentrations (0.01 -50 µM). After one hour of incubation at 37°C, medium containing drug was removed and fresh medium added. The plates were then incubated for 5 days before adding MTT to determine cell viability. Optical density was measured at 570 nm and melphalan zero concentration was considered as 100% viability. At concentration of 10 µM of melphalan, approximately 80% of DNA-PK deficient cells and 40% of DNA-PK proficient cells were eliminated.

### 4.8.2 Potential enrichment by melphalan

To study the possible enrichment of gene-repaired cells in polyclonal populations by melphalan treatment, those samples that showed positive gene repair results in the above *Bsa*WI assay (section 4.6) were again used in this experiment. These m*Tert scid* fibroblasts had been transduced (in triplicate as in section 2.2.14.2) with IDLVs encoding for SFFV ZFNs and *Prkdc-Bsa*WI template at a ratio of ZFN1: ZFN2:

Template (1:1:2) at MOI 500, 1000, 2500 or 5000. Mock (non-transduced), template only IDLVs (MOI 10000) and SFFV ZFN1 IDLVs (MOI 5000) transduced fibroblasts were used as negative controls. The samples were allowed to grow for 3 weeks post-transduction, with weekly treatments with 10 μM melphalan.

Melphalan belongs to the nitrogen mustard family of cross-linking anticancer drugs (De Silva *et al.*, 2000) that can induce DSB by a genotoxic mode not affecting cell survival at low concentrations (Vock *et al.*, 1999). Melphalan induced-DSBs can be repaired by NHEJ, and most crucially DNA-PK is determinant factor in the cellular response to this DNA damaging agent (Muller *et al.*, 2000). In a mixed cell population, DNA-PK proficient (gene-repaired) cells could be enriched with melphalan as they possess intact DNA repair mechanism whereas other DNA-PK deficient (*scid*) cells will be eliminated because of their faulty NHEJ DNA repair.

The results show (Figure 4.22) that the number of melphalan-resistant CFUs was remarkably vector dose-dependent: the higher the MOI used, the more melphalan resistant CFUs emerged. It is noteworthy that it was impossible to keep in culture the cells transduced with ZFN IPLVs, most likely because the continuous expression of the ZFNs in the cells may mediate cytotoxicity and cell loss.

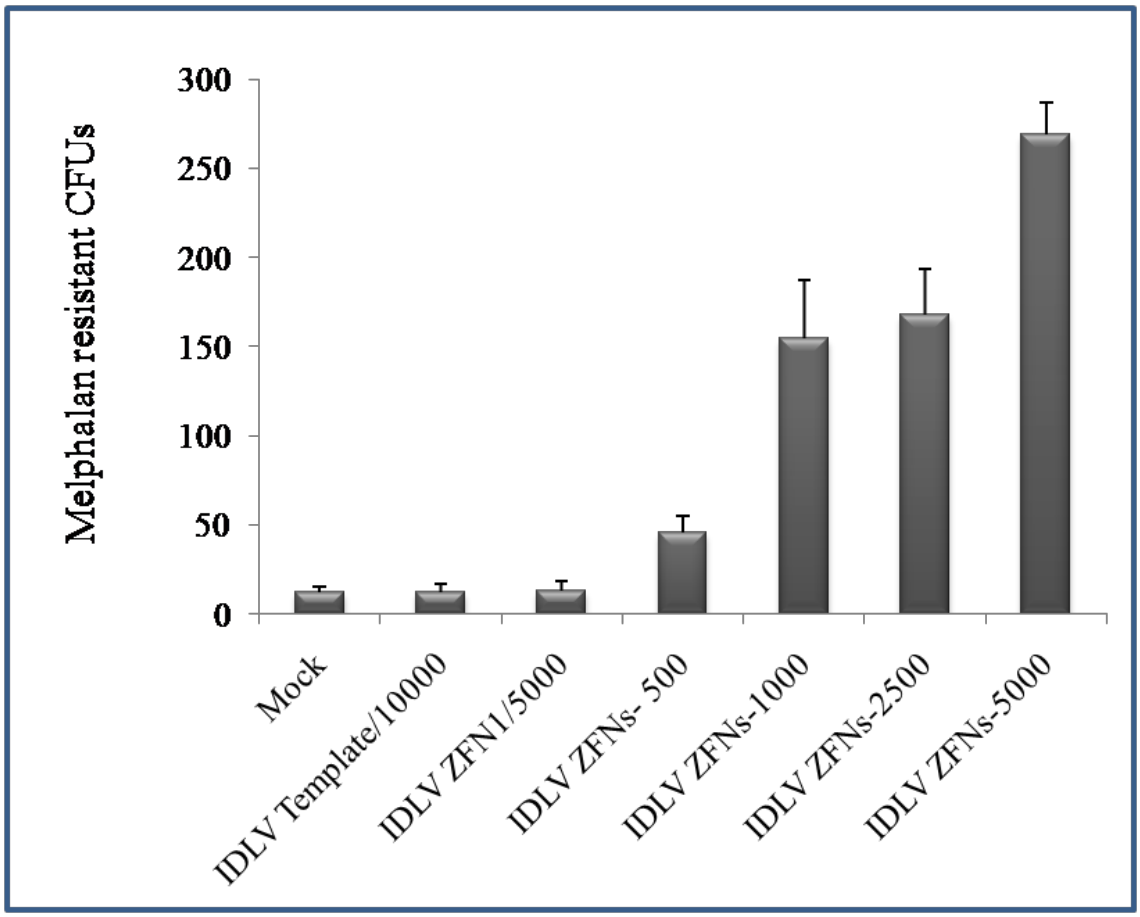



**Figure 4.22: Melphalan resistant CFUs after enrichment of gene repaired cells.** Gene-targeted cells with IDLVs encoding for SFFV ZFNs and corrective *Prkdc-Bsa*WI template were seeded in triplicate at the indicated MOIs. The next day, 10 μM of melphalan was added for 1 hour and then replaced by fresh medium. The cells were kept in culture for 3 weeks with weekly melphalan exposure, stained with crystal violet and finally CFUs were counted. The number of melphalan-resistant CFUs was significantly dose-dependent.

## 4.9 Detection of gene repair in melphalan enriched cells

To study whether melphalan-mediated enrichment in CFUs of the population of targeted m*Tert scid* fibroblasts from the previous experiment (section 4.8.2) has correlated with gene repair at the molecular level, the *Bsa*WI assay was repeated. For this, melphalan-enriched samples were subjected to this assay alongside to untreated gene-repaired cells.

Recapping, the m*Tert scid* fibroblasts were transduced in triplicate with IDLVs encoding SFFV-ZFNs and corrective *Prkdc-Bsa*WI template using a ratio of ZFN1: ZFN2: Template (1:1:2) at MOI 500 or 2500. Mock (non-transduced), *Prkdc-Bsa*WI template only IDLVs (MOI 10000) and SFFV ZFN1 IDLVs (MOI 5000)-transduced fibroblasts were used as negative controls. The positive control was prepared by PCR amplification of plasmid pHind-TK*Bsa*WI using the same primers that were used previously to amplify long probe of 1708 bp (Material and method section 2.2.12.3). The melphalan non-treated cells of controls and samples were harvested after 10 days and kept aside; while all melphalan-treated cells were maintained in culture for 3 weeks with weekly melphalan exposure. At three weeks post-transduction, melphalan resistant CFUs were counted (this part was completed above in section 4.8.2), then the CFUs were harvested and finally genomic DNA was extracted and subjected to the *Bsa*WI assay.

The results showed (Figure 4.23) that for both IDLV MOIs there was a considerable increase in gene repair frequency detected after enrichment with melphalan, compared to that without drug treatment. This suggests that gene-repaired cells have a selective advantage over uncorrected *scid* cells in the presence of melphalan.

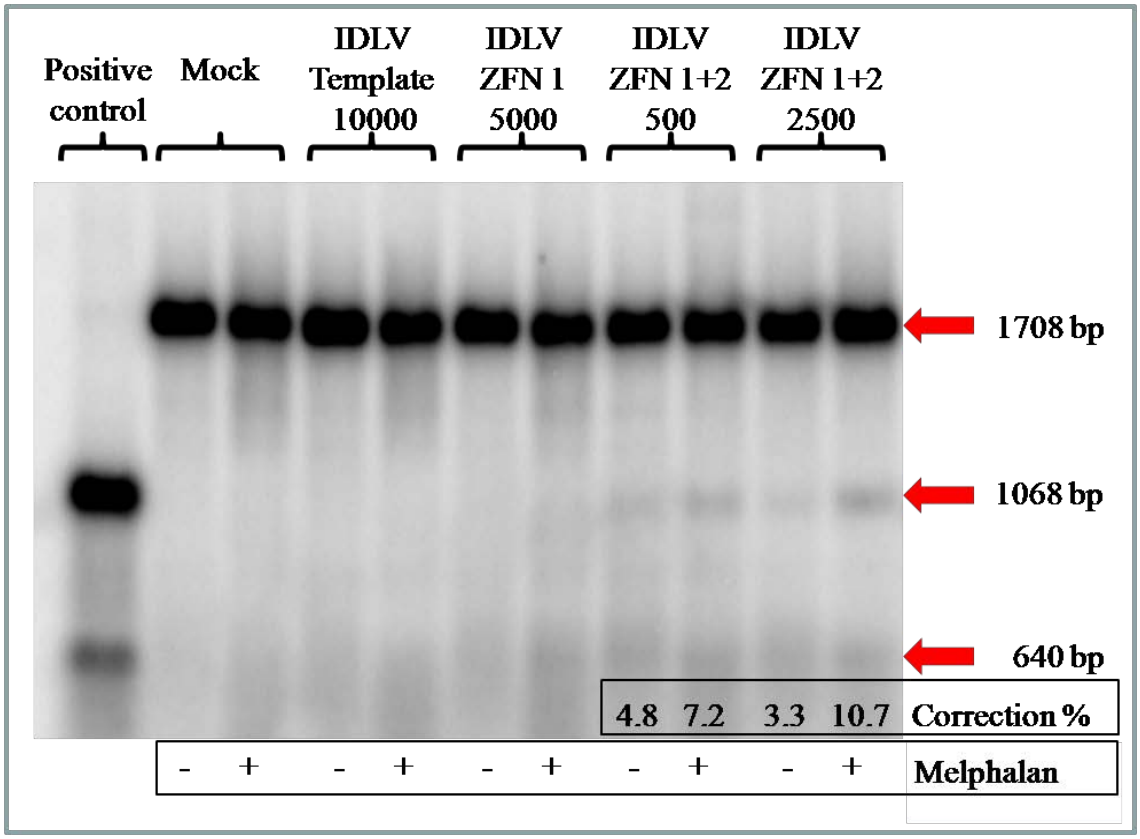



**Figure 4.23: Gene repair levels at *Prkdc* locus in polyclonal gene-repaired *scid* fibroblast populations, with or without melphalan enrichment.** The m*Tert scid* fibroblasts were transduced with IDLVs encoding the ZFNs and *Prkdc-Bsa*WI template at the indicated MOIs. The melphalan non-treated cells were harvested at ten days post-transduction, while other cells were subjected to weekly treatments of 10 μM melphalan for one hour, for 3 weeks. Finally genomic DNA was extracted and the *Bsa*WI assay was performed. The percentage of gene-repaired cells was determined as described earlier in main text using the small (1068-bp) probe. A considerable increase in gene repair frequency was detected after treatment with melphalan.

# Chapter Five

# ZFN activity in *scid* HSCs

# 5 ZFN activity in *scid* HSCs

In this part of work, I have optimised vectors and experiments to explore ZFN activity *ex vivo*. Two main objectives were planned for: (i) achieving the highest possible transduction efficiency in *scid* HSCs, as they are delicate to culture and hard to transduce; and (ii) demonstrating ZFN cutting by Surveyor *Cel*-I assay and then gene repair events by *Bsa*WI assay. To do this, I used wt HSCs to optimise HSC purification by magnetic cell depletion, transduced HSCs with eGFP lentiviral vectors, cultured the cells for three days and then assessed transduction efficiency and viability by flow cytometry. Having set up these parameters, I conducted working with *scid* HSCs, testing their viability and transduction with different lentiviral and adenoviral vectors driven by different promoters. Optimal parameters and vectors were then used to study ZFN cutting and gene repair.

## 5.1 Isolation of lin$^-$ HSCs

### 5.1.1 Isolation, depletion and transduction of wt lin$^-$ HSCs

Here I conducted a preliminary experiment to test the efficiency of isolation of wt lin$^-$ HSCs and to explore lentiviral transduction efficiency in those cells. The wt mice strain BALB/cOlaHsd was used to isolate wt bone marrow HSCs as described earlier (section 2.2.11). On the other hand, a lentiviral transfer plasmid pRRLsc-CEW was used to prepare IPLVs and IDLVs encoding for CMV-driven eGFP and WPRE element (CEW vectors). The aims of these experiments were to explore efficiency of isolation, transduction with CEW vectors and viability 72 hours post-transduction. Optimisation of these parameters will be helpful to set the major experiment with *scid* HSCs.

The average of total bone marrow isolated wt HSCs was ~ 4.8 X $10^7$ per 7-week old male mouse; whereas average of recovered lin$^-$ wt HSCs was ~ 6.5 X $10^5$ per mouse, ~ 1.4 % of original total cells. Lin$^-$ depleted population of wt HSCs were cultured at maximum density of 1 X$10^5$ cells/well and subsequently transduced with CEW IPLVs or IDLVs using qPCR MOIs 500, 1000 or 2000. After 72 hours the culture was terminated and eGFP expression and cell viability were examined by flow-cytometry. The viability was determined by DAPI staining at final concentration of 1µg/ml. In flow cytometric analysis loss of membrane integrity is an indicator of cell death; hence the

cells that exclude the dye are considered viable, while cells with non-intact membranes allow the dye inside the cell.

The results revealed (Figure 5.1) that the viability for the mock was 78 % and for other samples gradually decreased with increasing MOIs. eGFP transduction efficiency was 28 %, 41 % and 76 % when HSCs were transduced with CEW IPLVs at MOI 500, 1000 and 2000, respectively; it was 21 %, 35 % and 52 % when HSCs were transduced with CEW IDLVs at MOI 500, 1000 and 2000, respectively.

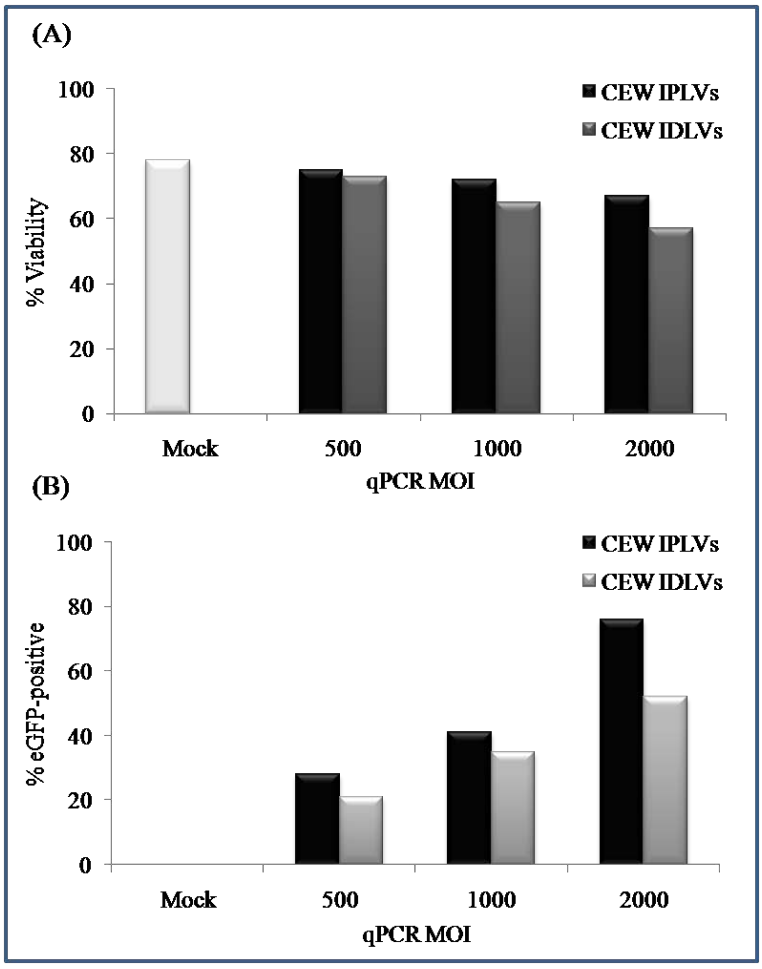



**Figure 5.1: Transduction of mouse wt lin$^-$ HSCs with CEW lentivectors.** Depleted lin$^-$ wt HSCs were transduced at the indicated qPCR MOIs with CEW IPLVs or IDLVs. (A) Viability (measured by DAPI staining) and (B) transduction efficiency were determined three days post-transduction by flow cytometry.

### 5.1.2 Isolation, depletion and transduction of *scid* lin$^-$ HSCs

The ultimate target cells to be corrected by gene repair in this project are the *scid* HSC progenitors, ideally using efficient non integrating vectors coding for ZFNs. Therefore; a series of experiments were carried out to explore the highest possible transduction efficiency that could be achieved in these cells. The *scid* mouse strain BALB/ cJHan (tm)Hsd-Prkdcscid was used to isolate *scid* lin$^-$ HSCs. The only known genetic alteration in these *scid* mice is the point mutation at *Prkdc* gene. The same protocol that was used to deplete and culture wt HSCs (section 5.1.1 above) was used again but this time using *scid* mice. These experiments aimed again at exploring: efficiency of isolation, transduction eGFP and viability 72 hours post-transduction.

For the aim of achieving the highest possible transduction in these cells, lentiviral vectors driven by different promoters or adenoviral vectors (serotype 5) driven by CMV promoter, all encoding an eGFP cassette, were used.

The average of total bone marrow isolated *scid* HSCs was ~ 2 X $10^7$ per 7-week old male mouse; whereas average of recovered lin$^-$ *scid* HSCs was ~ 5 X $10^5$ per mouse, ~2.5 % of original total cells. Lin$^-$ *scid* HSCs were cultured at same density and conditions used for wt HSCs. Initially I used CEW IPLVs and IDLVs to transduce lin$^-$ depleted population of *scid* HSCs using qPCR MOIs 500, 1000, 2500 or 5000. After 72 hours the culture was terminated and eGFP expression and viability were examined by flow-cytometry. The viability was determined by DAPI staining. The results obtained from this experiment (Figure 5.2) showed that the viability for the mock *scid* cells and IPLV-transduced HSCs was more than 80 % and about 70 % for the IDLV-transduced HSCs. Transduced cells had approximately the same viability levels regardless of MOI, suggesting that there were no toxic lentivector effects related to MOI increase. On the other hand, transduction efficiency of CEW IPLVs and IDLVs was similar and MOI dose-dependent, reaching more than 75 % of total population at MOI 5000. However; mean fluorescent intensity (MFI) index, which is an indicator of the cell's eGFP intensity, showed MOI-related increases eGFP for IPLVs but not for IDLVs.

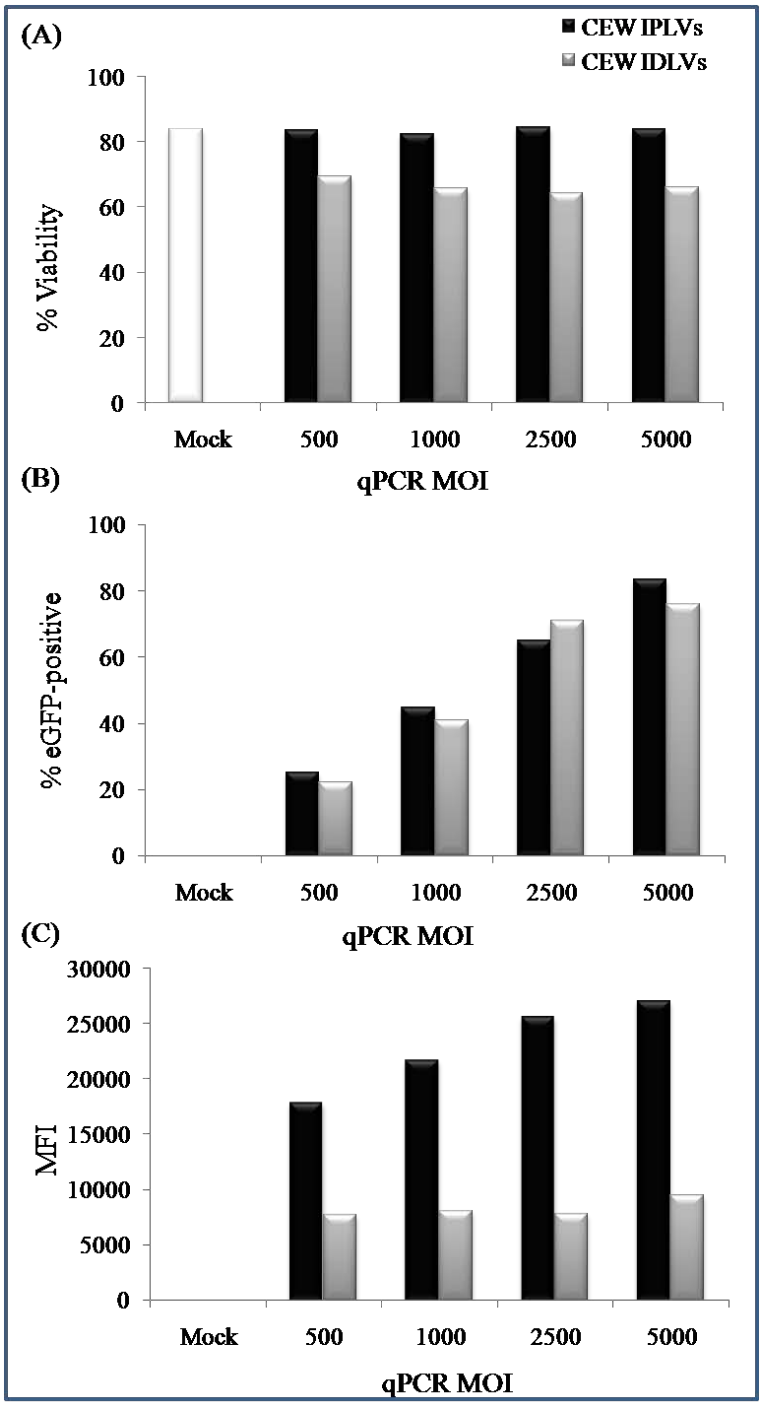



**Figure 5.2: Transduction of mouse *scid* HSCs with CEW lentivectors.** Depleted lin$^-$ *scid* HSCs were transduced at the indicated qPCR MOIs with CEW IPLVs or IDLVs. (A) Viability (measured by DAPI staining), (B) transduction efficiency and (C) MFI index were determined three days post-transduction by flow cytometry. Transduction efficiency was MOI dose-dependent for both vector types, but MFI index eGFP increased with dose only for IPLVs. There was no obvious vector-related toxicity.

Additionally, I used lentiviral transfer plasmid pHR'sc-SEW to prepare IPLVs and IDLVs encoding for SFFV-driven eGFP cassette and WPRE element (SEW vectors). Then I transduced $lin^-$ depleted population of *scid* HSCs with SEW IPLVs and IDLVs at qPCR MOI 500, 1000, 2500 or 5000. Three days later, I examined eGFP expression and viability (DAPI) by flow-cytometry.

The results (Figure 5.3) showed that the viability for the mock *scid* cells, IPLV- and IDLV-transduced HSCs was similar and above 75 %, suggesting that there were no lentivector-related toxic effects. Importantly, transduction efficiency of SEW IPLVs was MOI dose-dependent, reaching more than 90 % of total population at MOI 5000. Furthermore, the MFI index showed a dose-dependent increase eGFP with IPLVs. Yet SEW IDLVs transduction efficiency was very poor in these cells.

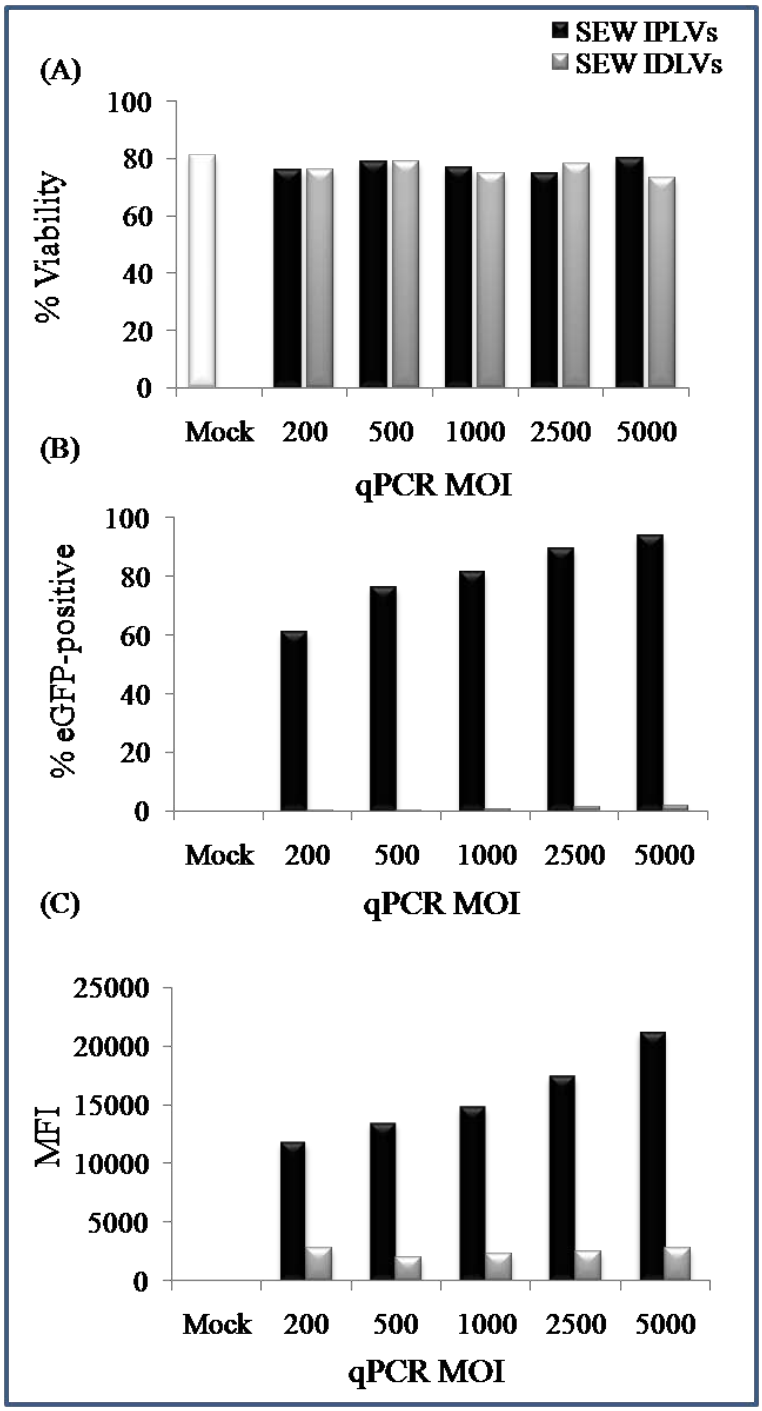



**Figure 5.3: Transduction of mouse *scid* HSCs with SEW lentivectors.** Depleted lin$^-$ *scid* HSCs were transduced at the indicated qPCR MOIs with SEW IPLVs or IDLVs. (A) Viability (measured by DAPI staining), (B) transduction efficiency and (C) MFI index were determined three days post-transduction by flow cytometry. Transduction efficiency and MFI index were MOI dose-dependent for IPLVs, but very poor with IDLVs eGFP. There was no vector-related toxicity.

In collaboration with Miss. Sara Oliván at RHUL a set of IDLVs encoding for eGFP but driven by different promoters was prepared. They included vectors prepared using the following lentiviral transfer plasmids (1) pRRLsc-eGFP-W to prepare control IDLVs coding for eGFP without external promoter, (2) pRRLsc-GFAP-eGFP-W to prepare IDLVs coding for eGFP driven by glial fibrillary acidic protein (GFAP) promoter, (3) pRRLsc-PGK-eGFP-W to prepare IDLVs coding for eGFP driven by phosphoglycerate kinase (PGK) gene promoter, (4) pRRLsc-SV40-eGFP-W to prepare IDLVs coding for eGFP driven by SV40 promoter, (5) pRRLsc-P11.5-eGFP-W to prepare IDLVs coding for eGFP driven by the African swine fever virus (ASFV) P11.5 promoter and (6) pRRLsc-P54-eGFP-W to prepare IDLVs coding for eGFP driven by the ASFV P54 promoter. Both P11.5 and P54 promoters are PCR-amplified from the ASFV genome. This virus is the etiological agent of a highly lethal hemorrhagic disease of domestic swine (Tulman *et al*., 2009).

The *scid* HSC lin$^-$ depleted population was then transduced with eGFP IDLVs controlled by different promoters at qPCR MOI 200, 1000 or 5000. Three days later, eGFP expression and viability (DAPI) were examined by flow cytometry. The results showed (see Figure 5.4) that the viability for the mock and IDLV-transduced cells was similar and higher than 70 %, suggesting that there were no lentiviral-related toxic effects. Unfortunately, none of the tested promoters led to efficient eGFP expression in lin- HSCs even at MOI 5000, the highest eGFP expression achieved being ~ 14 % with the P54 promoter followed by ~ 12 % with control IDLVs which could be explained to possible cryptic promoter active in HSCs in promoterless backbone, or residual activity of SIN LTR.

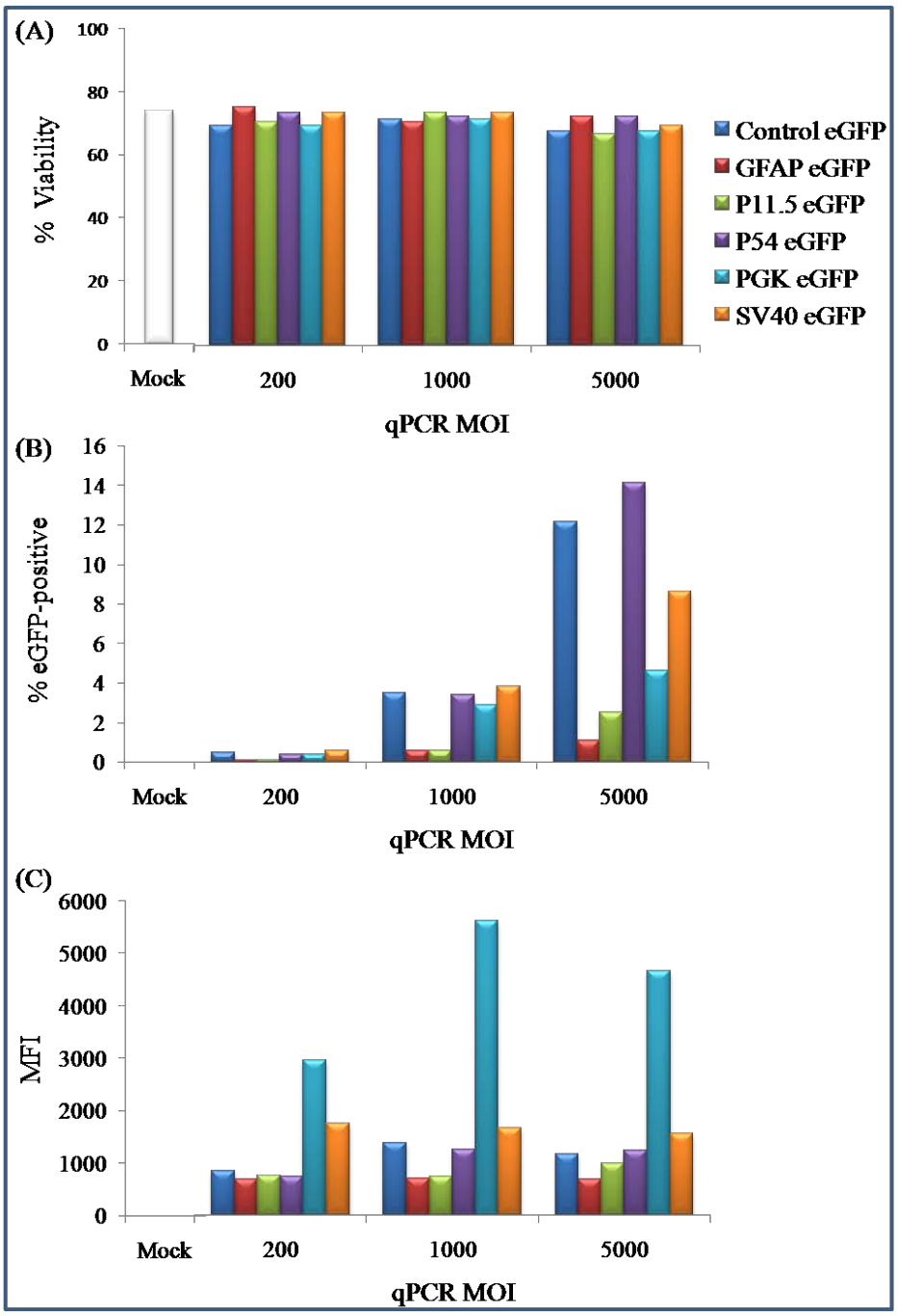



**Figure 5.4: Transduction of mouse *scid* HSCs with eGFP IDLVs driven by different promoters.** Depleted lin$^-$ *scid* HSCs were transduced at the indicated qPCR MOIs with eGFP IDLVs driven by either no external promoter (control), GFAP, P11.5, P54, PGK or SV40 promoters. (A) Viability (measured by DAPI staining), (B) transduction efficiency and (C) MFI index were determined three days post-transduction by flow cytometry. There was no vector-related cytotoxicity; yet the transduction efficiency was very poor.

Alternatively and for the same reason of exploring maximum possible transduction efficiency that could be obtained in these cells, three different adenoviral vectors (Ad 5) were used. They were a kind gift from Dr. Ramon Alemany, Institut Català d´Oncologia, IDIBELL, Spain. These adenoviral vectors were modified Ad5, expressing eGFP under the constitutive CMV promoter replacing E1 region. The mechanism of Ad5 cell infection involves two chronological virus–cell interactions; first binding of fiber knob domain to the primary CAR then followed by interaction of the Arg–Gly–Asp (RGD) motif of the penton base with cellular integrins initiating a faster adenovirus cell entry. The 3 received adenoviral vectors were: TL (AdTL) unmodified Ad 5, KKT (AdTKRGDK) modified adenovirus that has heparin sulfate glycosaminoglycans putative-binding site KKT inserted in the Ad5 fiber shaft domain, and TLRGD (AdTLRGD) modified adenovirus that has RGD motif inserted in the HI loop of the fiber knob. These insertions results in adenoviral enhanced cellular entry and different biodistribution profile (Alemany and Curiel, 2001; Bayo-Puxan *et al*., 2009; Sandovici *et al*., 2006). For gene therapy applications, adenovirus serotype 5 could be good vectors as they can be produced at high titres and efficiently transduce a broad range of cell types (Bayo-Puxan *et al*., 2006).

This experiment was done by transduction of *scid* HSCs $lin^-$ depleted population with the said eGFP adenoviral vectors controlled by CMV promoter at qPCR MOI 10000, 25000, 50000, 100000, 200000 or 400000. Then after three days, eGFP expression and viability (DAPI) were examined by flow-cytometry. The results revealed (Figure 5.5) that the viability for the mock *scid* HSCs, AdTL and AdTKRGDK-transduced cells was more than 70 %. However; viability of AdTLRGD-transduced cells decreased particularly with increasing MOI, suggesting that there was a related cytotoxic effect. This experiment showed clearly that while the MFI achieved could be relatively high, the percentage of eGFP expression was very poor even at a very high MOI of 400000. The highest eGFP expression, noticed with AdTLRGD adenoviral vectors, was correlated with reduced viability. It can be concluded from this experiment that none of the tested adenoviral vectors expressed eGFP efficiently in these cells and they were not investigated further. The main conclusion that can be made from these experiments is that among the tested vectors, configurations and promoters, SFFV IPLVs and CMV IDLVs were the most efficient.

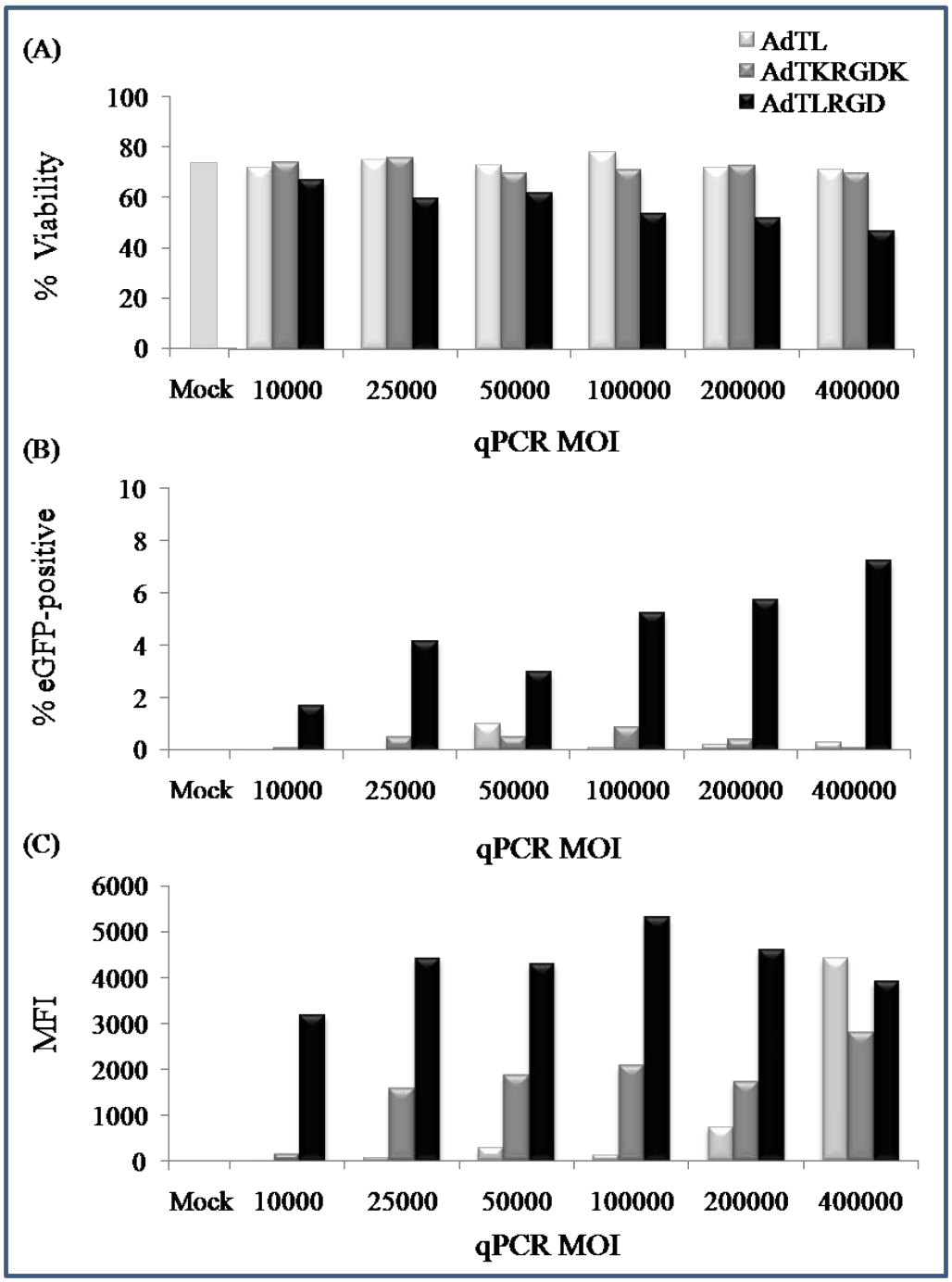



**Figure 5.5: Transduction of mouse *scid* HSCs with CMV eGFP adenoviral vectors.**

Purified lin$^-$ *scid* HSCs were transduced at the indicated qPCR MOIs with CMV eGFP adenoviral vectors. (A) Viability (measured by DAPI staining), (B) transduction efficiency and (C) MFI index were determined three days post-transduction by flow cytometry. Transduction with AdTL and AdTKRGDK did not cause cytotoxicity; in contrast to AdTLRGD. Clearly, adenoviral transduction efficiency was very poor.

## 5.2 Testing the efficiency of ZFN cutting in *scid* HSCs

I have shown in the previous chapter that ZFN-induced-DSBs repaired by NHEJ in *scid* fibroblasts can be detected by the Surveyor *Cel*-I mutation detection assay; likewise I used the same assay to examine ZFN-induced-DSBs in *scid* HSCs. The Surveyor *Cel*-I assay, described earlier in section 2.2.9, can detect mismatches in heteroduplex DNA following target site repair by NHEJ.

Lin$^-$ *scid* HSCs were purified and transduced with either CMV or SFFV ZFN IPLVs or IDLVs. IPLV transduction was done with qPCR MOI 200, 500, 1000 or 2500, and for IDLVs qPCR MOI 200, 500, 1000, 2500 or 5000 was used. Three days post-transduction, the samples were collected and viability was determined by flow cytometry using DAPI staining. Genomic DNA was extracted and then subjected to Surveyor *Cel*-I assay. Same as in fibroblasts, PCR amplification generates a 304-bp parental fragment (for DNA sequence please see Appendix 1, Box 1) and upon cutting by Surveyor's nuclease (at ZFN repaired DSB) two smaller fragments of 184 and 126 bp are generated. Samples were run in PAGE, visualized under UV light, images taken and finally cutting percentages were calculated using ImageJ.

The results shown in Figure 5.6 demonstrate efficient ZFN cutting (and hence expression) when SFFV IPLVs were used, while nothing could be detected by this assay in the other samples tested (SFFV IDLVs, CMV IPLVs and IDLVs). Furthermore, cell viability (see Figure 5.7 ) was strongly affected proportionally to Surveyor ZFN cutting with SFFV IPLVs. Additionally, those samples that were treated with CMV IDLVs had lower viability than those treated with SFFV IDLVs and CMV IPLVs samples; this could be ascribed to the fact that the CMV IDLV titre was lower than that of others, requiring a higher volume of viral suspension for transduction to obtain the desired MIOs. Adding a higher volume of viral suspension would dilute the HSC medium, perhaps causing a reduction in viability.

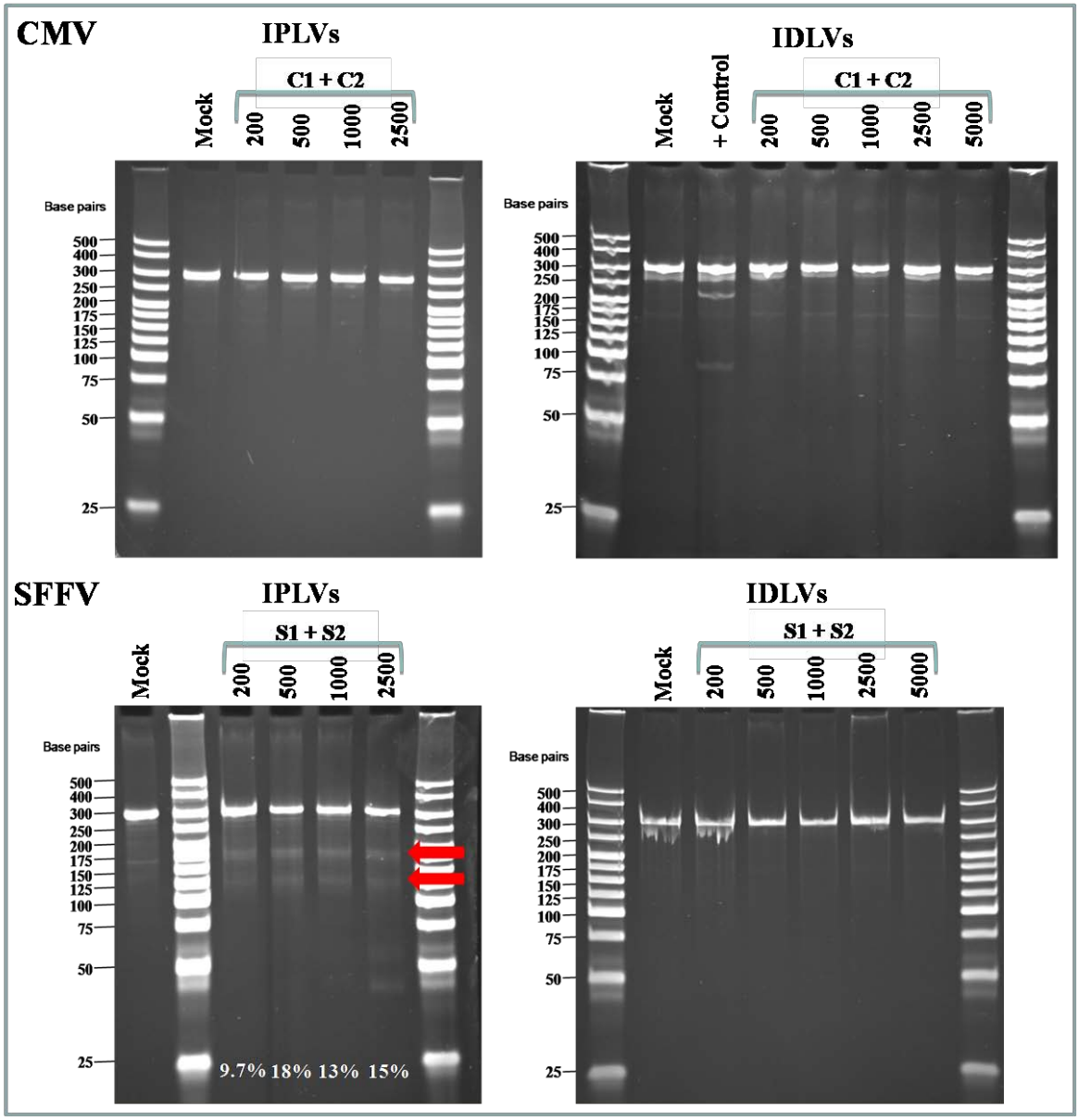

CMV
IPLVs
C1 + C2
Mock
200
500
1000
2500
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25
IDLVs
Mock
+ Control
C1 + C2
200
500
1000
2500
5000
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25
SFFV
IPLVs
S1 + S2
Mock
200
500
1000
2500
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25
9.7% 18% 13% 15%
IDLVs
S1 + S2
Mock
200
500
1000
2500
5000
Base pairs
500
400
300
250
200
175
150
125
100
75
50
25

**Figure 5.6: Site-specific cutting by ZFNs determined by Surveyor *Cel*-I assay in *scid* HSCs.** Purified lin$^-$ *scid* HSCs were transduced at the indicated qPCR MOIs with either CMV ZFN lentivectors (upper panel) or SFFV lentivectors (lower panel). Three days later cells were harvested, genomic DNA was extracted and then samples were examined by Surveyor *Cel*-I assay. The PCR amplification generates a 304-bp parental fragment and upon Surveyor's nuclease cutting (at ZFN target) two smaller fragments of 184 and 126 bp (red arrows) are produced and visualised in 10 % PAGE. The results confirmed that there was efficient ZFN expression and cutting with SFFV IPLVs but not with other combinations. Percentages represent ZFN cutting percentage calculated using ImageJ. M: mock, C1: CMV ZFN1, C2: CMV ZFN2, S1: SFFV ZFN1 and S2: SFFV ZFN2. + Control was a mixture of scid/wt fibroblasts (50/50) % used as a positive control; Surveyor's nuclease cut at *scid* point mutation site generates fragments of 88 and 216 bp.

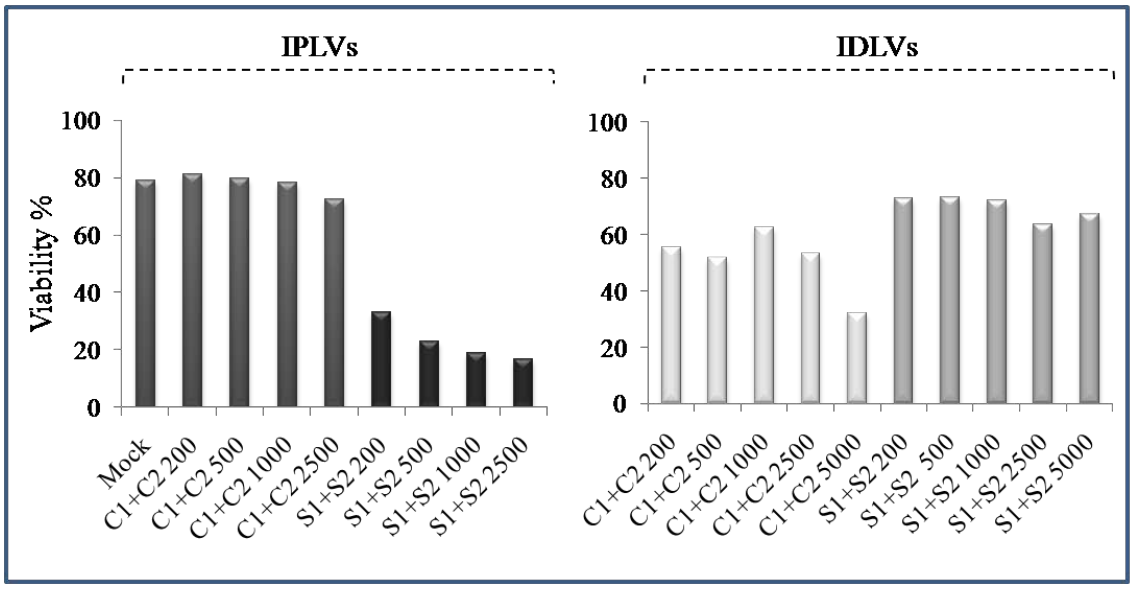



**Figure 5.7: Viability of *scid* HSCs transduced with either CMV or SFFV ZFN IPLVs and IDLVs.** *Scid* HSCs were transduced at the indicated MOIs with ZFN IPLVs or IDLVs driven by either CMV or SFFV. Viability was determined three days later by flow cytometry using DAPI staining. Cell viability was strongly affected in a manner proportional to ZFN cutting when cells were transduced with SFFV IPLVs. C1: CMV ZFN1, C2: CMV ZFN2, S1:SFFV ZFN1 and S2: SFFV ZFN2.

I have shown earlier in section 4.5.1 that the minimal percentage of mismatches that can be detected by Surveyor *Cel*-I assay was 5 %. In order to improve the sensitivity of InDel detection in samples that did not show clear cutting in the Surveyor *Cel*-I assay, a selected set of samples (SFFV IDLV-treated) was subjected to deep sequencing. This was done in collaboration with Dr. Manfred Schmidt at National Center for Tumor Diseases, Heidelberg, Germany. Depending on the depth of coverage, each base can be read hundreds of times during the sequencing process allowing identifying of very rare sequence mutations (Morozova *et al*., 2009).

The PCR amplicons of 304-bp resulting from transduction of *scid* HSCs with SFFV IDLVs at qPCR MOI 200, 500, 1000, 2500 or 5000 were shipped to complete this analysis. The results in Table 5.1 revealed that ZFNs driven by SFFV in non integrating lentivector configuration can induce low but detectable levels of DSBs that can be

repaired by NHEJ and measured as InDel events. The highest % of InDels was observed at MOI 2500, being approximately 0.5 %.

**Table 5.1: Deep sequencing results in SFFV IDLV-treated *scid* HSCs**

| Sample | Sequenced molecules | InDels | InDels % |
|---|---|---|---|
| Mock | 15163 | 12 | 0.08% |
| SFFV IDLVs (S1+S2) 200 | 4214 | 4 | 0.09% |
| SFFV IDLVs (S1+S2) 500 | 5723 | 15 | 0.26% |
| SFFV IDLVs (S1+S2) 1000 | 12832 | 45 | 0.35% |
| SFFV IDLVs (S1+S2) 2500 | 18262 | 90 | 0.49% |
| SFFV IDLVs (S1+S2) 5000 | 11234 | 52 | 0.46% |

## 5.3 Detection of gene repair by *Bsa*WI assay in *scid* HSCs

The *Bsa*WI assay was used again in this experiment so that HR events in targeted *scid* HSCs could be detected exploiting the silent diagnostic restriction site for *Bsa*WI in the *Prkdc-Bsa*WI corrective template. During the process of gene repair, the supplemented corrective template containing the diagnostic *Bsa*WI site would mediate the transfer of the site to the host genome. Eventually the presence of the *Bsa*WI site in *Prkdc-Bsa*WI template would be diagnostic of gene repair occurrence. This work was done in collaboration with Dr. Céline Rocca at RHUL.

The *Bsa*WI assay was done as described earlier (section 2.2.12); in brief the lin$^-$ *scid* HSCs were targeted with lentivectors encoding the *Prkdc-Bsa*WI corrective template of 1626 bp and either CMV or SFFV ZFNs. For CMV or SFFV ZFN IPLVs three vectors

were used for cell co-transduction: two ZFN IPLVs (ZFN1 and ZFN2) and *Prkdc-Bsa*WI template IDLVs with a ratio of ZFN1: ZFN2: template (1: 1: 2). For CMV or SFFV IDLVs two vectors were used for transduction, where the template and each of the ZFN monomers were encoded by the same vector, keeping ratios of ZFN1: ZFN2: template (1: 1: 2). A non-transduced (mock) sample was used as a negative control, while the plasmid pHind-TK*Bsa*WI (containing a 9210 bp fragment of *Prkdc* around the *scid* site and with the *Bsa*WI target sequence introduced by site-directed mutagenesis) was used as a template to amplify the positive control by PCR.

Three days post-transduction, genomic DNA was extracted and the sequence around the *scid* target site was PCR amplified using primers external to the *Prkdc-Bsa*WI corrective template. Next, the amplicons were digested with *Bsa*WI to assay for the presence of the diagnostic feature (*Bsa*WI site). Later, the transferred blot was hybridised with the long radiolabelled probe (1708 bp), scanned and imaged. Finally, gene repair percentages were determined using ImageJ by quantification of band intensities (i.e. 1708 and 1068) then the following formula was applied:

Gene repair % = $(100 \text{ X } 1.6 \text{ X } I_{1068}) / (I_{1708} + [1.6 \text{ X } I_{1068}])$.

The results (see Figure 5.8) revealed that a faint band of gene repair could be detected only when SFFV ZFN IPLVs were used at MOI 100, with detection of the 1068 bp being diagnostic for gene repair. No detectable gene repair could be observed with SFFV IDLVs or CMV IPLVs and IDLVs.

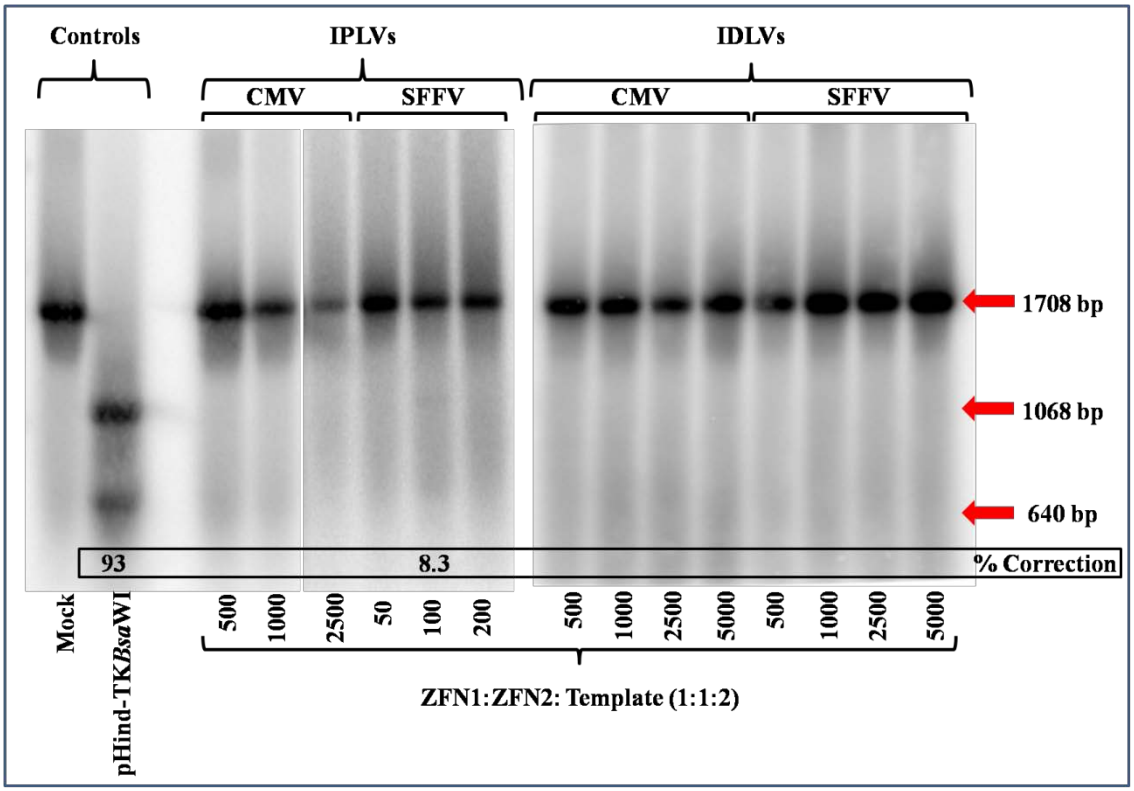



**Figure 5.8: Gene repair at *Prkdc* locus in *scid* HSCs, measured by *Bsa*WI assay.** Three days post-transduction, the *Bsa*WI assay was used to detect gene repair events. The presence of the smaller 1068-bp band was diagnostic for gene repair. Gene repair percentages were determined using ImageJ by quantification of band intensities as described in main text. Gene repair could be detected only when SFFV ZFN IPLVs were used at MOI 100. Controls: mock were non-transduced fibroblasts and pHind-TK*Bsa*WI plasmid containing the template modified by site-directed mutagenesis. IPLV samples: *scid* HSCs were transduced at the indicated MOIs with IPLVs driven by either CMV or SFFV. Three vectors were used: ZFN1 and ZFN2 IPLVs, and template IDLVs, at a ratio of ZFN1: ZFN2: template (1: 1: 2). For example, the number 500 on the figure indicates that the cells were transduced with MOI 500 of each ZFN vector and MOI 1000 of *Prkdc-Bsa*WI template vector. IDLV samples: *scid* HSCs were transduced at the indicated MOIs with IDLVs driven by either CMV or SFFV. Two vectors were used: template/ZFN1 and template/ ZFN2 IDLVs, at a ratio of ZFN1: ZFN2: template (1: 1: 2). For example the number 2500 on the figure indicates that the cells were transduced with MOI 2500 of each of template/ZFN1 and template/ ZFN2 IDLVs.

# Chapter Six

# *Ex vivo* gene repair and transplantation

# 6 *Ex vivo* gene repair and transplantation

Here I conducted the main experiment in the current study; the *ex vivo* gene repair of *scid* HSCs and transplantation after optimisation of *ex vivo* protocol. The recipient mice were divided into four groups of five animals each. I carried out the purification of lin$^-$ HSCs then I transduced them overnight with lentiviral vectors where applicable according to animal groups. Next I did the transplantation of potentially corrected *scid* HSCs into irradiated *scid* recipients. The recipient mice were bled at weeks 9, 16 and 24 post-transplantation. The multicolour flow cytometric analysis was performed for all animals by targeting specific CD markers in B cells, T cells and other haematopoietic cells. The transplantation experiment was completed with help of Dr. Michael Blundell; while the multicolour flow cytometric analysis was done with help of Dr. Maria Alonso-Ferrero.

## 6.1 Transplantation of lin$^-$ HSCs

In the previous chapters I developed a new gene therapy strategy based on gene repair using ZFN technology and IDLV platform delivery to rescue the classical *scid* mouse (*Prkdc* point mutant). Importantly, I have already demonstrated by *Cel*-I assay and deep sequencing in *scid* fibroblasts and HSCs that the ZFNs cut specifically their target site. Furthermore; I used ZFN- and template-mediated gene targeting to correct the *Prkdc* point mutation via the incorporation of "*neo*" selection cassette and/or a diagnostic feature of *Bsa*WI restriction site from the donor templates into the targeted locus. Also I showed restoration of DNA-PKcs functionality and increased resistance to DNA damage upon gene repair in *scid* fibroblasts; while in *scid* HSCs, gene repair was confirmed only when the ZFN genes were delivered by IPLVs. Considering all above promising finding, I proceeded to the *ex vivo* gene repair transplantation of corrected *scid* HSCs into *scid* recipients to attempt immune system reconstitution.

### 6.1.1 Transplantation experiment design

In this transplantation experiment, I designed 4 different groups contained 5 animals each: wt control, eGFP control, ZFN IPLVs and ZFN IDLVs (see Table 6.1). The donors were either wt male BALB/c OlaHsd mice for the wt control (group 1) or *scid*

male BALB/c JHan(tm)Hsd-Prkdcscid mice for the eGFP control (group 2) and ZFN groups (3 and 4). I used *scid* female BALB/c JHan(tm)Hsd-Prkdcscid mice as recipients (Table 6.1). The donor mice were euthanized by inhalation of a 75 % CO2 / 25 % O2 mixture and hematopoietic cells were harvested from the bone marrow of femurs and tibia then the lineage cell depletion kit was used as usual for lin$^-$ HSCs purification as described before (section 2.2.11).

**Table 6.1: Animal groups used in transplantation experiment.**

| Group # | Group name | Donors (males) | qPCR MOI | Recipients (females) |
|---|---|---|---|---|
| 1 | Wt control | (BALB/c OlaHsd) | - | BALB/c JHan(tm)Hsd-Prkdcscid |
| 2 | eGFP control | BALB/c JHan(tm)Hsd-Prkdcscid transduced with CEW IPLVs | 400 | BALB/c JHan(tm)Hsd-Prkdcscid |
| 3 | ZFN IPLV | BALB/c JHan(tm)Hsd-Prkdcscid transduced with template IDLVs, SFFV IPLVs ZFN1 and ZFN2 | 200: 100: 100 | BALB/c JHan(tm)Hsd-Prkdcscid |
| 4 | ZFN IDLV | BALB/c JHan(tm)Hsd-Prkdcscid transduced with IDLVs template/CMV ZFN1 and template/CMV ZFN2 | 500: 500 | BALB/c JHan(tm)Hsd-Prkdcscid |

Then the *scid* lin$^-$ HSCs in eGFP control group were transduced overnight (see next section) with CEW IPLVs at 400 qPCR MOI to match the total MOI that will be used for ZFN IPLVs group. In parallel the sample of ZFN IPLV group were transduced overnight with *Prkdc-Bsa*WI template IDLVs, SFFV ZFN1 and ZFN2 IPLVs at qPCR MOI 200: 100: 100, respectively. This combination is similar to that one I used earlier in (section 5.3) which demonstrated gene repair of 8.3% in *scid* HSCs. Simultaneously the sample of ZFN IDLV group were transduced overnight with *Prkdc-Bsa*WI

template/CMV ZFN1 and *Prkdc-Bsa*WI template/CMV ZFN2 IDLVs at qPCR MOI 500: 500; respectively.

### 6.1.2 Overnight transduction and preparation for injection

Following transduction, the lin$^-$ HSCs cells of wt control group and other transduced *scid* lin$^-$ HSCs for other groups were cultured at $10^6$ cells/ml in Stemspan medium containing 100 ng/ml Murine SCF, 100 ng/ml Murine Flt-3, 20 ng/ml human IL-6 and 1 % PenStrep at 5 % CO2, 37 °C for 16 hours. Next day the HSCs were prepared for intravenous injection into irradiated recipient mice; briefly the culture medium was removed and then the cells were washed and resuspended in PBS. All the recipients were sublethally irradiated with a total of 2.5 Gy/mouse in a single dose same day before injections. As these *scid* animals were *Prkdc* deficient and lacking efficient DNA damage response, I used a lower irradiation dose than that used earlier for other *scid* types (Fulop and Phillips, 1986; Mostoslavsky *et al*., 2006).

### 6.1.3 Transplantation

The recipient mice were used at 7 weeks of age and were maintained in specified pathogen free conditions. Each recipient was injected with 1 X $10^6$ cells in 200 µl plain PBS using intravenous injection route into the tail vein by 27 G bevel needle attached to a tuberculin syringe. Other gene therapy strategies used to rescue other *scid* models have used lower numbers of injected cells than in this study as they were simply based on gene addition approach (Huston *et al*., 2011); here the transplantation was conducted with maximum experimentally feasible HSC numbers in order to maximise the chances of transplanting corrected cells into the recipients.

Using DAPI staining and flow cytometry, the viability of all cells to be transplanted was checked out one hour before their injection. Additionally, the majority of the cells in eGFP control group were transplanted back into the recipients; however one well of $10^5$ cells of each of mock and eGFP IPLV-transduced *scid* lin$^-$ HSCs were kept in culture for 72 hours post-transduction to subsequently test their eGFP expression by flow cytometry.

The transplantation was done for two groups per day in two successive days: the control groups in the first day and the ZFN groups the next day. Flow cytometry results for viability of lin$^-$ HSCs before injection of all groups are shown in Figure 6.1. For wt lin$^-$ HSCs (wt control group) and *scid* mock or eGFP transduced (eGFP control group) the viability was about 45 %; while for ZFN groups was about 70 %. As in the current study the viability was normally close to 70 %, the lower viability values observed for the two control groups were likely due to technical issues noticed during lin$^-$ HSC depletion.

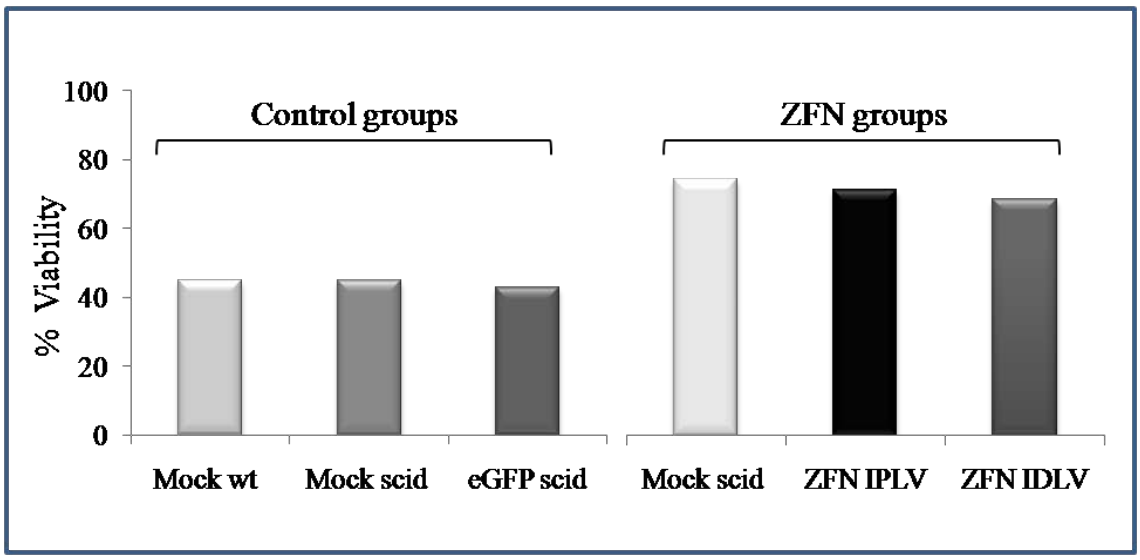



**Figure 6.1: Viability of mouse lin$^-$ HSCs one hour before transplantation.** Purified wt and/or *scid* lin$^-$ HSCs of transplantation groups (wt control, eGFP control, ZFN IPLVs and ZFN IDLVs) were transduced where applicable with lentiviral vectors and cultured at $10^6$ cells/ml for 16 hours. The viability was measured by DAPI staining followed by flow cytometry. The results showed that viability for the control samples (three leftmost columns) was about 45 %; while for the other groups was about 70 %.

On the other hand, flow cytometry results for eGFP expression of mock and eGFP IPLV-transduced *scid* lin$^-$ HSCs of eGFP group at 72 hours post-transduction are shown in Figure 6.2. The results demonstrated that the eGFP expression of eGFP IPLV-transduced *scid* lin$^-$ HSCs was 28 %, this finding will be informative as it will be used as a reference to compare with *in vivo* eGFP expression in the myeloid lineage subsequently.

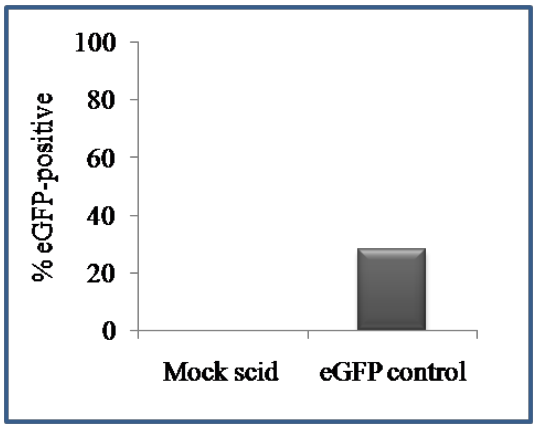



**Figure 6.2: Transduction of *scid* lin$^-$ HSCs (eGFP control group) with CEW IPLVs.** Purified *scid* lin$^-$HSCs were transduced CEW IPLVs at qPCR MOI 400. Next 72 hours the culture was terminated and eGFP expression was determined by flow-cytometry. The results revealed that eGFP expression was 28 %.

## 6.2 *Scid* phenotype correction.

### 6.2.1 Animal bleeding

In order to evaluate the engraftment and the immune system restoration in the transplanted animals by tracking immune cell reconstitution, all animals were bled at intervals of 9, 16 and 24 weeks post-transplantation. Samples of ~50 µl of freshly isolated peripheral blood from the tail vein were collected in EDTA tubes using a 27 G bevel needle attached to a tuberculin syringe. The samples were then prepared as described in (section 2.2.15.2).

#### 6.2.1.1 Week 9 bleeding

Multicolour flow cytometric analysis of all samples, to test for gene-repaired *scid* lin$^-$ HSC engraftment, was done by staining each blood sample with three different antibodies. The first was anti-human/mouse CD45R (B220) PE antibody to target B-lymphocytes. The CD45R is a 220-kDa cell-surface glycoprotein expressed mainly by the B-cell lineage from early pro-B to mature B cells. CD45R, along with other CD45 isoforms, helps to control different cellular functions like cell growth, differentiation and mitotic cycle (Monteith *et al*., 1996; Paulovicova *et al*., 2010).

The second antibody was anti-mouse CD3e PerCP-Cy5.5, to target T-lymphocytes. The CD3e (e for epsilon) is a 20-kDa subunit of the T-cell receptor (TCR) complex, required together with other CD3 subunits, gamma and delta, for appropriate assembly, trafficking and surface expression of the TCR complex (Fernandes *et al*., 2012). CD3e expression correlates with T cell maturation, thus it is expressed by thymocytes (T cells precursor) in a developmentally regulated manner and by all mature mouse T cells (Lai *et al*., 1998).

The third antibody was anti-mouse CD11b APC to target mainly myeloid cells. It is a commonly used marker for the monocyte/macrophage lineage, natural killer (NK) cells and granulocytes (Lai *et al*., 1998). CD11b is a 170 kDa glycoprotein non-covalently associated with CD18 to form The CD11b/CD18 integrin which helps to control inflammation by accelerating the apoptotic elimination of extravasated neutrophils (Castaigne *et al*., 2002).

Additionally anti-mouse Ig, k /negative control (FBS) compensation beads were used (section 2.2.15.3) along with the other blood samples as negative control. The compensation was set using rat IgG2a K isotype control PE, Armenian hamster IgG isotype control PerCP-Cy5.5, rat IgG2b K isotype control APC or all of them together. Similarly anti-human/mouse CD45R (B220) PE antibody, anti-mouse CD3e PerCP-Cy5.5 antibody or anti-mouse CD11b APC were used with control beads for gate compensation.

Flow cytometric results revealed (Figure 6.3 and 6.4) that all *scid* recipient animals transplanted with wt HSCs were fully reconstituted, matching B and T cell levels in normal animals. These findings indicated that the transplantation approach was successful and all preparation steps before and during transplantation experiment were proficiently performed. In the group of *scid* animals transplanted with eGFP-transduced *scid* cells there was no reconstitution other than in the myeloid lineage, and the average of eGFP expression in peripheral blood mononuclear cells (PBMCs) was ~ 32 %. The presence of this control was very important as it showed the background levels of both B and T cells in non-rescued animals which will be helpful to compare with ZFN-treated groups. At the same time, eGFP detected fluorescence will give an indication of

myeloid engraftment efficiency of lentiviral vector-transduced HSCs. Another finding in this group was that the average of total detected myeloid cells for the five animals was ~ 54 %, representing the main blood cell population.

For groups ZFN IPLVs and ZFN IDLVs, there were no major increases observed in B and T cell levels. However animal number 2 and animal number 3 of ZFN IPLV and ZFN IDLV groups, respectively, showed promising results. The animal number 2 of ZFN IPLV group showed ~ 4.5 % of T cells, whereas animal number 3 of ZFN IDLV group showed ~ 1.5 % of T cells. Crucially, even more important than the percentage values was the observed appearance of the T-cell populations (Figure 6.3). Both observed populations were clearly isolated and well-shaped pointing out that they might be real rescued populations.

On the other hand, no significant B-cell compartments were observed in either ZFN IPLV or ZFN IDLV groups. A previous study of murine *scid*-X1 (**γ** chain) model rescue has suggested that reconstitution of B cells is slower than that of T-cells and requires higher numbers of transplanted cells (Huston *et al*., 2011). As the correction strategy attempted in my study is based on gene repair, it is expected that slow reconstitution would be normal as a result of the number of rescued cells being low. My observation of detectable levels of T-cells but not of B-cells in ZFN-treated groups correlates with the published work.

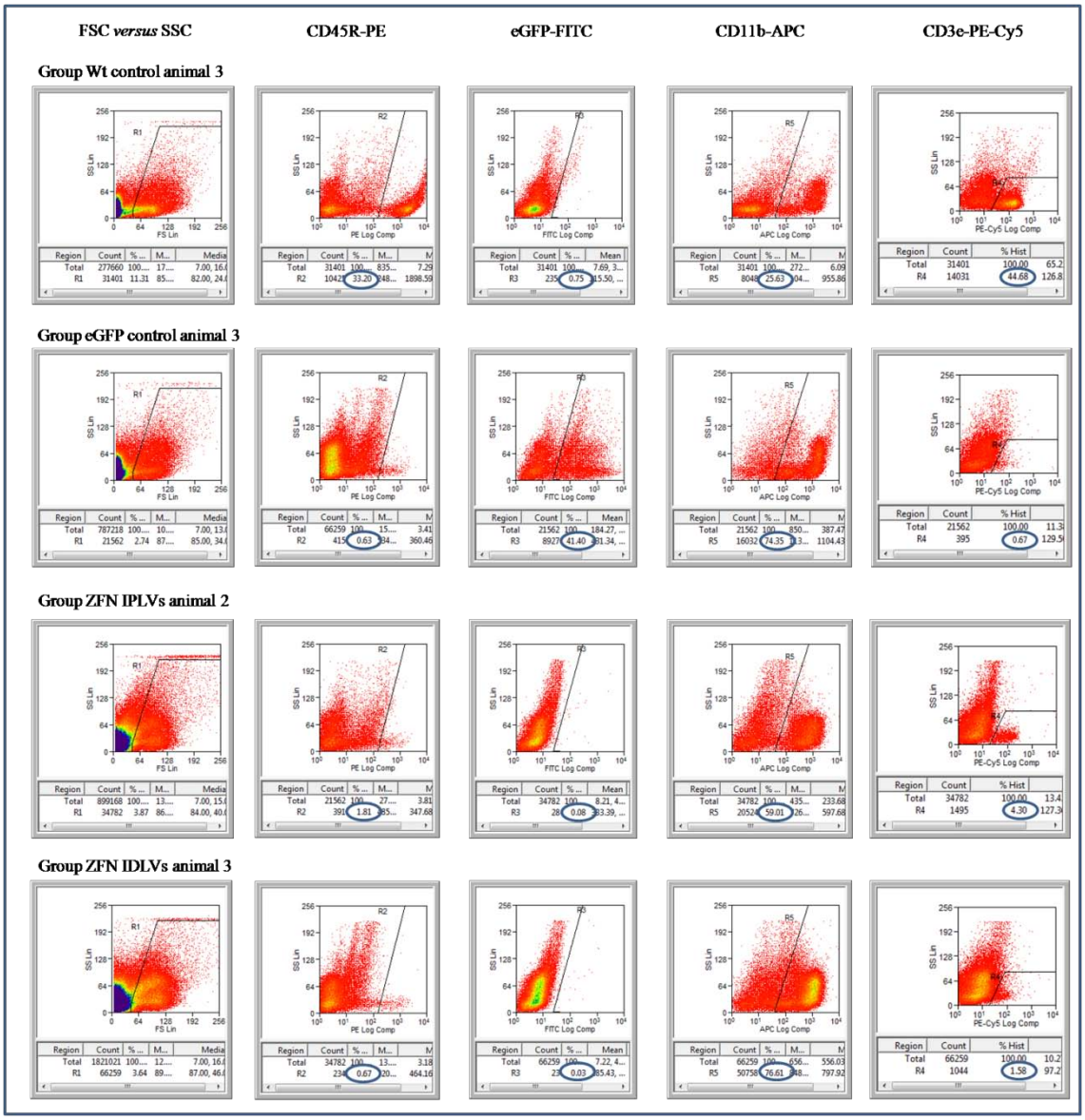



**Figure 6.3: Flow cytometry analyses of peripheral blood from selected samples at 9 weeks post-transplantation.** Blood samples were collected from recipient mice at 9 weeks post-transplantation and multicolour flow cytometric analysis of HSC engraftment was completed for all animals by staining for CD45R, CD3e and CD11b. Images of representative animals of each control group and selected animals from treated groups are showed here. The animals of the wt control group were fully reconstituted, while only myeloid reconstitution was clearly observed in the eGFP control group (eGFP-positive cells are also CD11b-positive, not shown). However one animal from each of ZFN IPLV and ZFN IDLV groups showed promising increased levels of T cells.

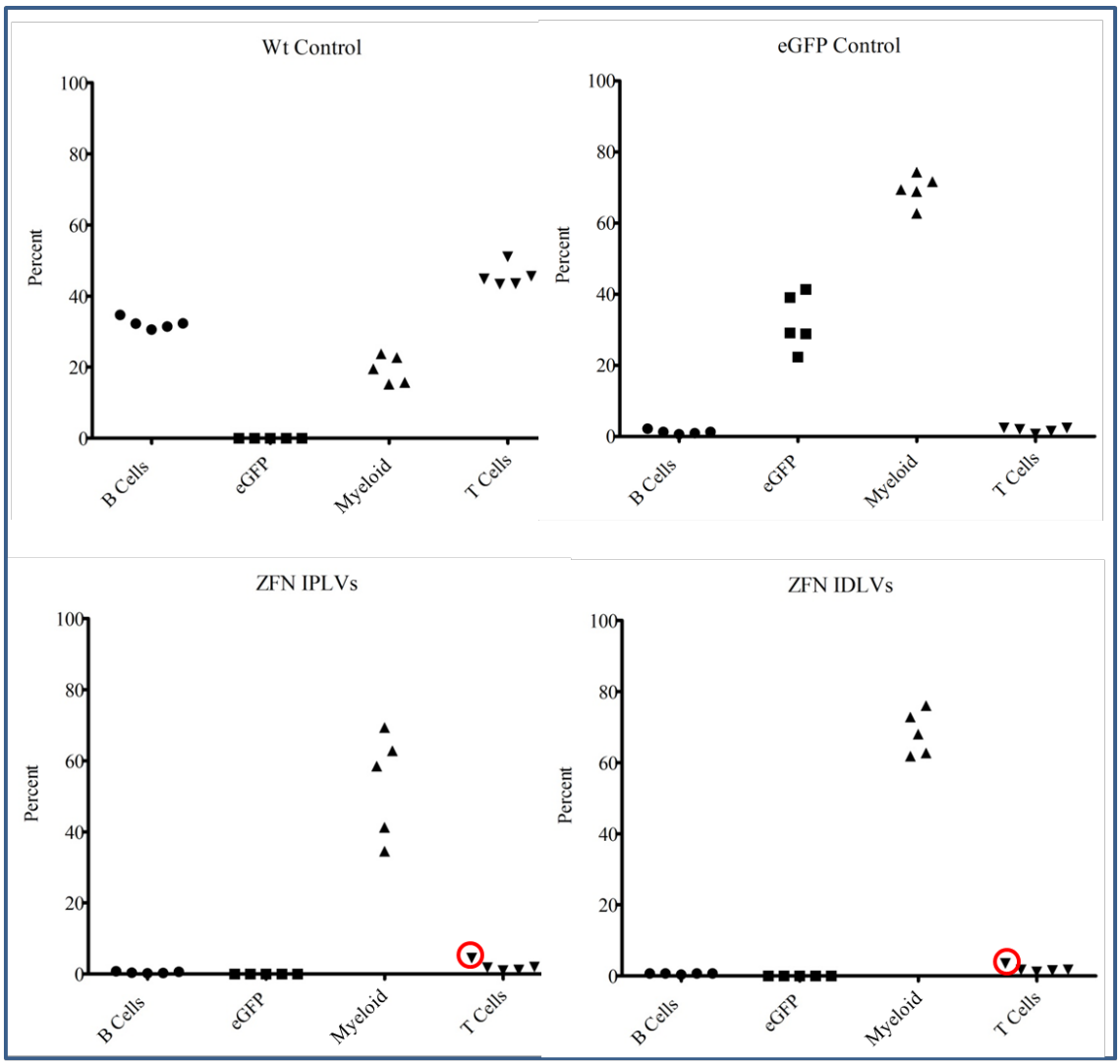



**Figure 6.4: Percentages of CD45R$^+$, eGFP$^+$, CD3e$^+$ and CD11b$^+$ cells in peripheral blood of all animals at 9 weeks post-transplantation.** Recipient mice were subjected to 2.5 Gy irradiation and transplanted with either 1 × $10^6$ wt lin$^-$ cells or *scid* lin$^-$ cells transduced with eGFP IPLVs, ZFN IPLVs or ZFN IDLVs. At week 9 post-transplantation, animals were bled and multicolour flow cytometric analysis of engrafted HSCs was conducted for all animals by staining PBMCs for CD45R, CD3e and CD11b. All animals in group wt control were convincingly fully reconstituted demonstrating normal levels of B and T cells; while none were in eGFP control, ZFN IPLV and ZFN IDLV groups. However one animal in each group of ZFN IPLV and ZFN IDLV showed promising increased levels of T cells (red circles).

#### 6.2.1.2 Week 16 bleeding

As in week 9 above, all samples were stained with antibodies targeting CD45R (B220), CD3e and CD11b. Additionally, flow cytometric analysis for gene-repaired *scid* lin$^-$ HSCs engraftment was carried out using two additional antibodies to target mature T cells in selected samples of each animal group. The first was anti-Mouse CD4 PE to target helper T cells, a subset of mature T cells. The CD4 is a 55 kDa cell surface marker could be found in a majority of thymocytes, helper T cells and dendritic cells. This receptor plays a vital role both in T cell development and in optimal functioning of mature T cells as it is initiates the early phase of T-cell activation (Lai *et al*., 1998; Yucel *et al*., 2004).

And the second: anti-Mouse CD8a APC to target cytotoxic T cells, a subset of mature T cells. The CD8a is an approximately 34 kDa cell surface marker expressed by a majority of thymocytes and a subpopulation of mature TCR T cells. CD8 plays a role in T cell development and activation of mature T cells by mediating T cell killing effector (Lai *et al*., 1998; Maldonado-Lopez *et al*., 1999).

Simultaneously an age-matched wt female mouse of strain BALB/c OlaHsd was bled and used as a control by staining with antibodies against CD45R, CD3e and CD11b. As in week 9, compensation was set using rat IgG2a K isotype control PE, Armenian hamster IgG isotype control PerCP-Cy5.5 rat IgG2b K isotype control APC or all of them together.

The results of flow cytometry confirmed again (as illustrated in Figure 6.5 and 6.6) that all recipient animals in wt control group were fully reconstituted and matched B and T cell levels in normal animals. Once more these findings proved that the transplantation experiment and procedure were successful. Meanwhile flow cytometric results of eGFP control group demonstrated that the average of eGFP expression was ~ 37 % and there was no reconstitution other than in myeloid lineage, as expected in this control group. Another finding in this group was that the average of total detected myeloid cells for the five animals was ~ 70 % representing the major population of blood cells.

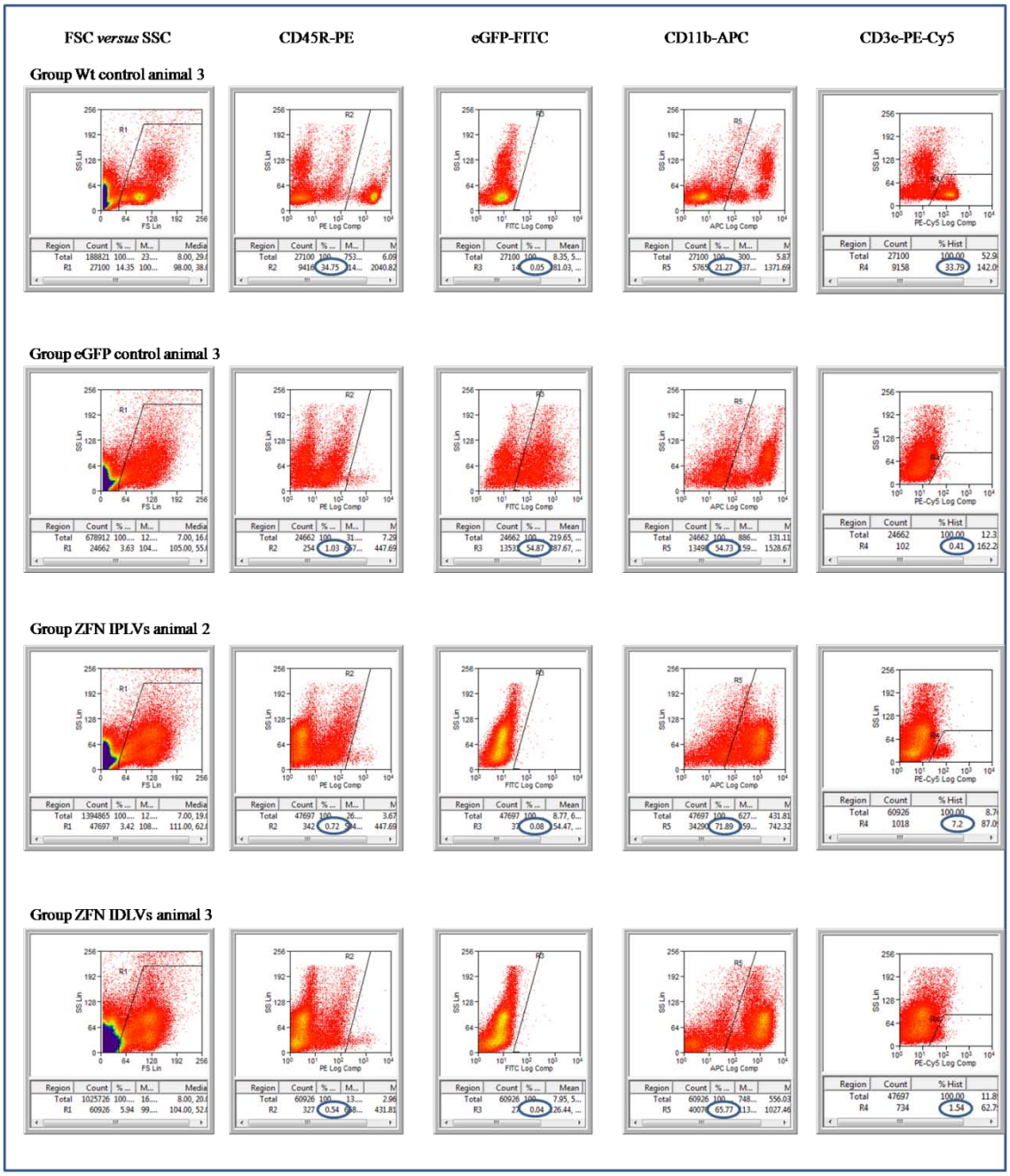



**Figure 6.5: Flow cytometry analyses of peripheral blood in selected samples at 16 weeks post-transplantation.** Blood samples were collected from recipient mice at 16 weeks post-transplantation and multicolour flow cytometric analysis for HSC engraftment was completed in all animals by staining for CD45R, CD3e and CD11b. Images of representative control animals and selected ZFN animals are shown here. Animals of wt control group were fully reconstituted; while that was not the case in eGFP control group. Confirming the results from the 9-week bleed, one animal in each of ZFN IPLV and ZFN IDLV groups showed promising increased levels of T cells.

For groups ZFN IPLV and ZFN IDLV, there were no major increases observed in B and T cell levels. However animal number 2 and animal number 3 of ZFN IPLV and ZFN IDLV groups, respectively, demonstrate again promising results. Animal number 2 of ZFN IPLV group showed even higher (~ 7 %) T-cell levels than that in previous bleeding, whereas animal number 3 of ZFN IDLVs group showed ~ 1.5 % of T cells. Nevertheless the main finding here was that the appearance of the detected T cell populations, isolated and well shaped, again suggests that these two populations were real rescued populations (Figure 6.5). On the other hand, no increased levels were detected in B cell compartments in either ZFN IPLV and ZFN IDLV groups, as expected.

In order to further explore whether those increased levels of T cell populations observed in the two animals of ZFN IPLV and ZFN IDLV groups represented mature and functional T cells, PBMCs from these animals were stained with anti CD4 and CD8a. The results obtained from multicolour flow cytometric analysis are presented in Figure 6.7. In the representative animal from wt control group, the results indicated that within the T-cell population there were two subpopulations belonging to mature T-cell subsets. They were ~ 24 % $CD4^+$ and ~ 6 % $CD8a^+$ T cells, representing the frequencies of normal BALB/c mouse peripheral blood. Conversely, in a representative animal from eGFP control group there were ~ 0.25 % $CD4^+$ and ~ 0.13 % $CD8a^+$ T cells, representing the observed reference of non-rescued animals.

The clearer evidence observed in this experiment was with animal 2 from ZFN IPLV group, in which ~ 3.1 % $CD4^+$ and ~ 0.62 % $CD8a^+$ T cells were detected. Furthermore the results showed that these T-cell subpopulations were well isolated and compact, suggesting that they are real and potentially functional T cell populations. Similarly with animal 3 from ZFN IDLV group there were ~ 1.57 % $CD4^+$ and ~ 0.5 % $CD8a^+$ T-cells. The overall conclusion from week 16 post transplantation bleeding suggested that there was an emergence of rescued, mature and possibly functional T cell populations, as an indicator of slow immune system reconstitution.

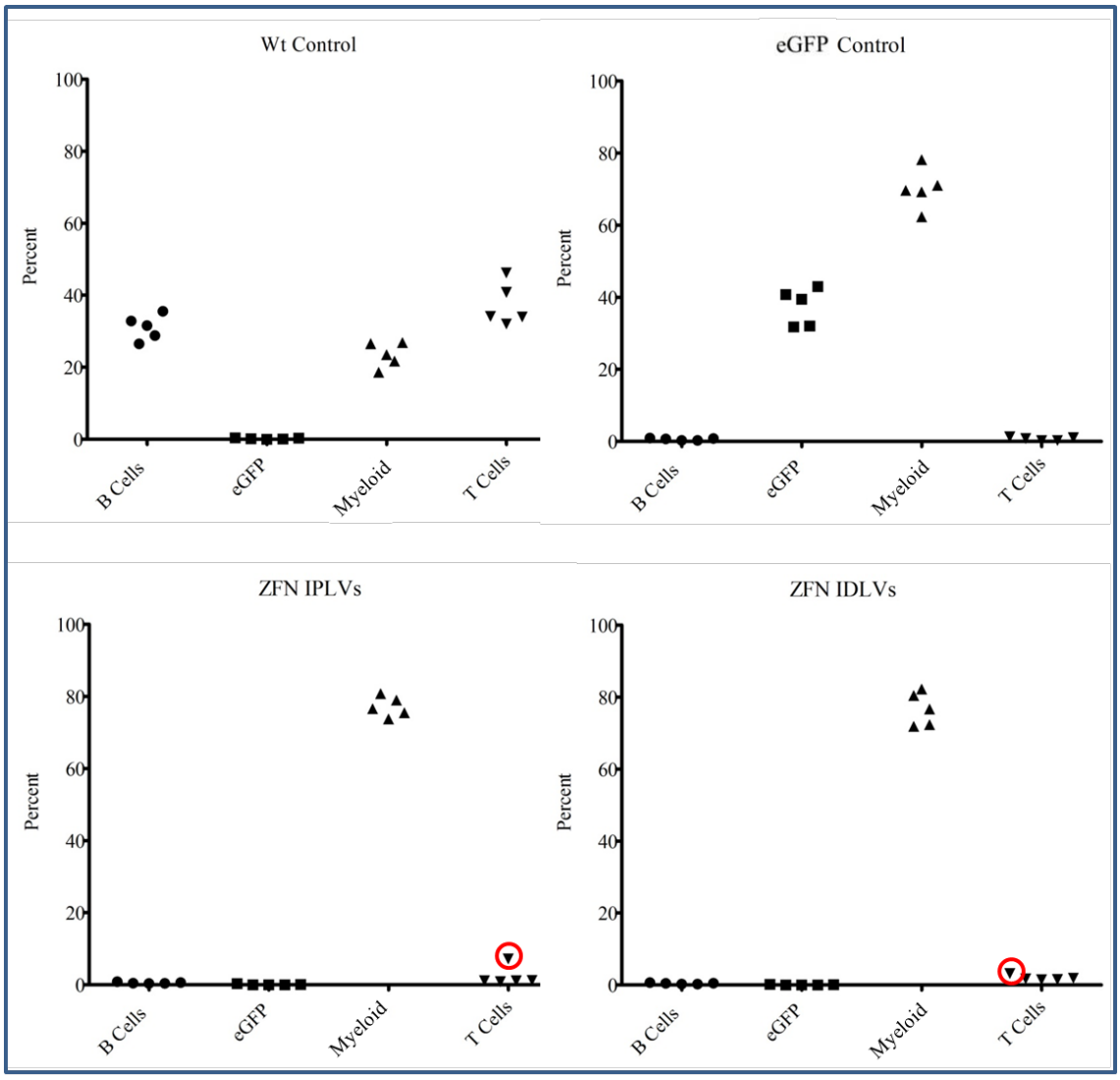



**Figure 6.6: Percentages of CD45R$^+$, eGFP$^+$, CD3e$^+$ and CD11b$^+$ cells in peripheral blood of all animals at 16 weeks post-transplantation.** Recipient mice were subjected to 2.5 Gy irradiation and transplanted with either 1 × $10^6$ wt lin$^-$ cells or *scid* lin$^-$ cells transduced with eGFP IPLVs, ZFN IPLVs or ZFN IDLVs. At week 16 post-transplantation, animals were bled and multicolour flow cytometric analysis for HSC engraftment was done for all animals by staining for CD45R, CD3e and CD11b. All animals in group wt control were fully reconstituted; while none were in eGFP control group. In ZFN IPLV and ZFN IDLV groups those two animals that showed increased levels of T cells in previous bleeding had even higher levels of T cells this time (red circles).

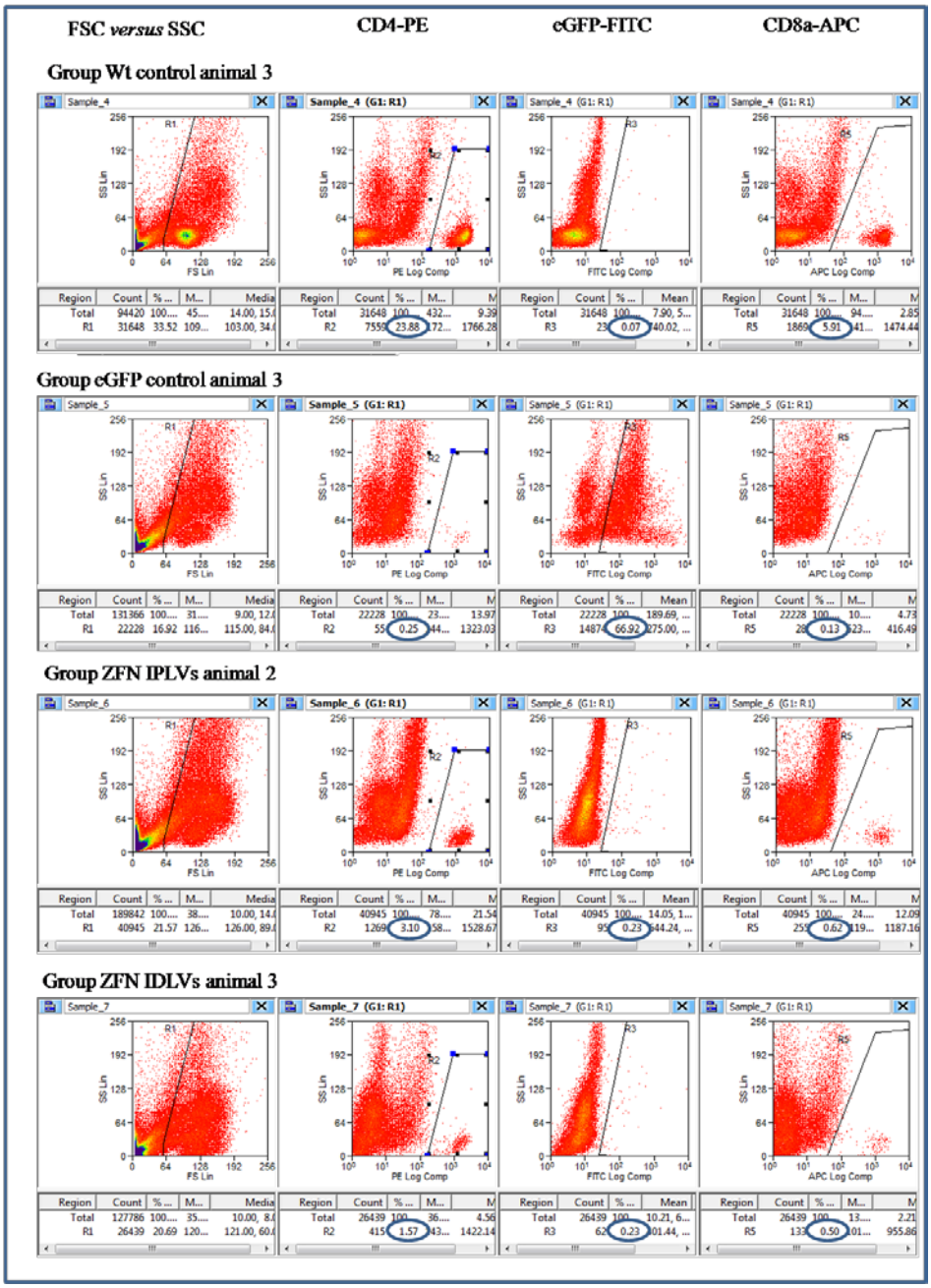



**Figure 6.7: Flow cytometry analysis of CD4 and CD8 analyses in peripheral blood from selected samples at 16 weeks post-transplantation.** Blood samples were collected from recipient mice at 16 weeks post-transplantation and multicolour flow cytometric analysis for HSC engraftment was completed for one animal of each group by staining with anti CD4 and CD8a. The animal from wt control group showed ~ 24 % $CD4^+$ and ~ 6 % $CD8a^+$, the eGFP control showed ~ 0.25 % $CD4^+$ and ~ 0.13 % $CD8a^+$, responder from ZFN IPLV group showed ~ 3.1 % $CD4^+$ and ~ 0.62 % $CD8a^+$ and that from ZFN IDLV group showed ~ 1.57 % $CD4^+$ and ~ 0.5 % $CD8a^+$ T cells.

#### 6.2.1.3 Week 24 bleeding

At 24 weeks post-transplantation, all animals were again bled and stained with antibodies targeting CD45R (B220), CD3e and CD11b in addition to matured T cell markers CD4 and CD8a as explained earlier. Age-matched wt female mice of strains BALB/c OlaHsd and BALB/c JHan(tm)Hsd-Prkdcscid were used as positive and negative controls, respectively. The compensation was set as before with rat IgG2a K isotype control PE, Armenian hamster IgG isotype control PerCP-Cy5.5 rat IgG2b K isotype control APC or all of them together.

Flow cytometry results are shown in Figure 6.8 , 6.9, and 6.10. Altogether, the results confirmed again that all recipient animals in group wt control were fully reconstituted and matched B and T cell levels in normal animals, proving that the transplantation experiment and procedure were successful. Yet again flow cytometric results of eGFP control group showed that the eGFP expression average was ~ 27 % and there was no detected reconstitution beyond the myeloid compartment (as expected) in this control group. Another finding in this group was that the average of total detected myeloid cells for all animals in this group was ~ 75 % as the major population of blood cells.

The responder animal from ZFN IPLV group (animal number 2) showed T cell levels even higher than previously observed, reaching ~ 12 % T cell, compared to the T-cell average of wt control group of ~ 45 %. This finding represents major evidence of T cell reconstitution in this *scid* animal. On the other hand, animal number 3 of group ZFN IDLV demonstrated again an increased level of T cells at ~ 2 %. In both animals, well isolated T cell populations and clear T cell level increases (see Figure 6.8) suggest that immune reconstitution has begun, albeit at different speed. In contrast, no increased levels were detected yet in B cell compartments in either ZFN IPLV or ZFN IDLV groups, most likely because gene repair strategy would lead to slow reconstitution in general and even slower for B cells.

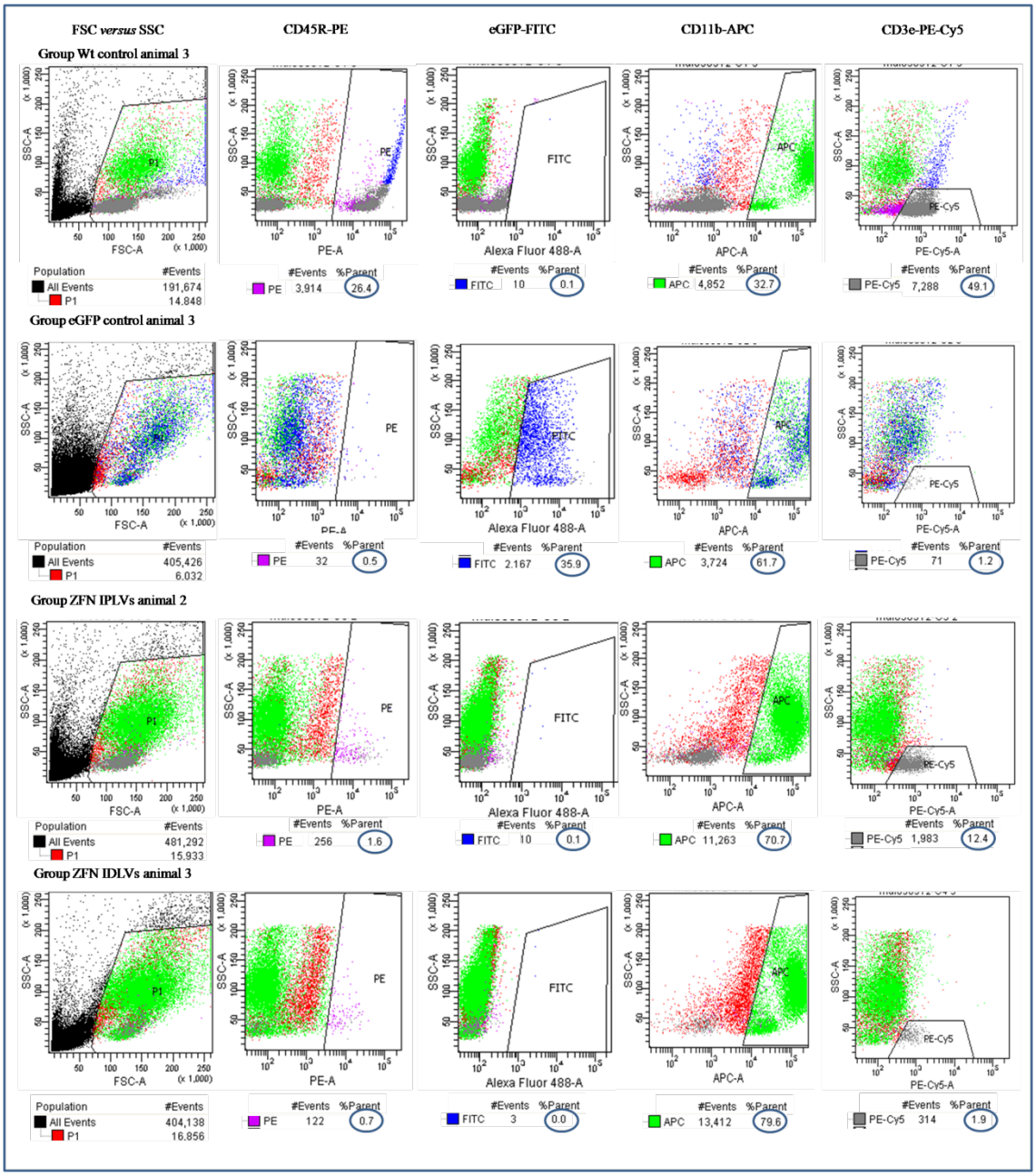



**Figure 6.8: Flow cytometry analysis of peripheral blood from selected samples at 24 weeks post-transplantation.** Blood samples were collected from recipient mice at 24 weeks post-transplantation and multicolour flow cytometric analysis for HSC engraftment was completed for all animals by staining for CD45R, CD3e and CD11b. Images of one animal of each group are shown here. The animal of group wt control was fully reconstituted and representative of the group; no reconstitution beyond myeloid compartment was observed in eGFP control. However, the responder animals from groups ZFN IPLV and ZFN IDLV showed increased and well-isolated T cell populations.

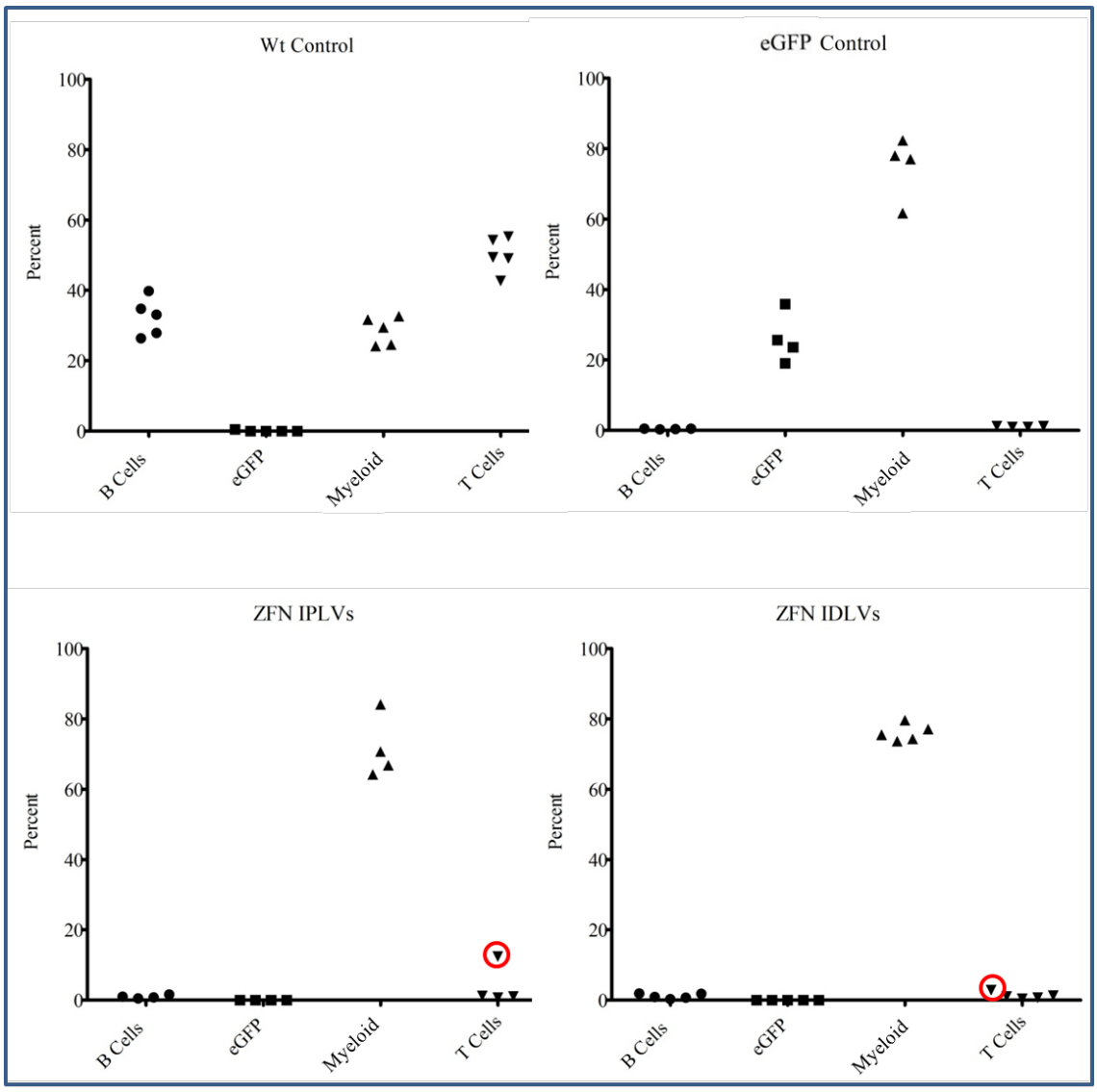



**Figure 6.9: Percentages of CD45R$^+$, eGFP$^+$, CD3e$^+$ and CD11b$^+$ cells in peripheral blood of all animals at 24 weeks post-transplantation.** Recipient mice were subjected to 2.5 Gy irradiation and transplanted with either 1 × $10^6$ wt lin$^-$ cells or *scid* lin$^-$ cells transduced with eGFP IPLVs, ZFN IPLVs or ZFN IDLVs. At week 24 post-transplantation, animals were bled and multicolour flow cytometric analysis for HSC engraftment was done for all animals by staining with CD45R, CD3e and CD11b antibodies. All animals in group wt control were fully reconstituted, while none were in eGFP control group. The number 2 animal of group ZFN IPLV showed distinctly increased level of T cells, while number 3 animal of ZFN IDLV group remains at an increased but steady T-cell level. Those two animals are circled in red.

To confirm that the T-cell populations from the responder ZFN mice contained mature T-cell subpopulation as an indication of restoration of functional immune cells, PBMCs from representative controls and the responder animal from each ZFN group were stained with anti CD4 and CD8a. Multicolour flow cytometric analysis results are presented in Figure 6.10. The animal from wt control group showed two subpopulations of functional T cell subsets. There were ~ 26 % of $CD4^+$ and ~ 5.5 % of $CD8a^+$ T cells, similar to normal BALB/c mouse peripheral blood. In contrast the animal from the eGFP control group had ~ 0.5 % $CD4^+$ and ~ 0.3 % $CD8a^+$ T cells, indicative of the reference value of non-rescued animals.

Evidence confirming restoration of mature immune cells was obtained from animal number 2 of ZFN IPLV group, which had ~ 10 % $CD4^+$ and ~ 3 % $CD8a^+$ T cells, observed as well-isolated T cell subsets. Animal number 3 of ZFN IDLV group showed ~ 2.3 % $CD4^+$ and ~ 0.8 % $CD8a^+$ T cells, suggesting again the presence of two mature T cell subsets. It can be concluded that at week 24 post-transplantation there was clear evidence of rescued, mature T cell population in two responder mice, as an indicator of a possible slow immune system reconstitution.

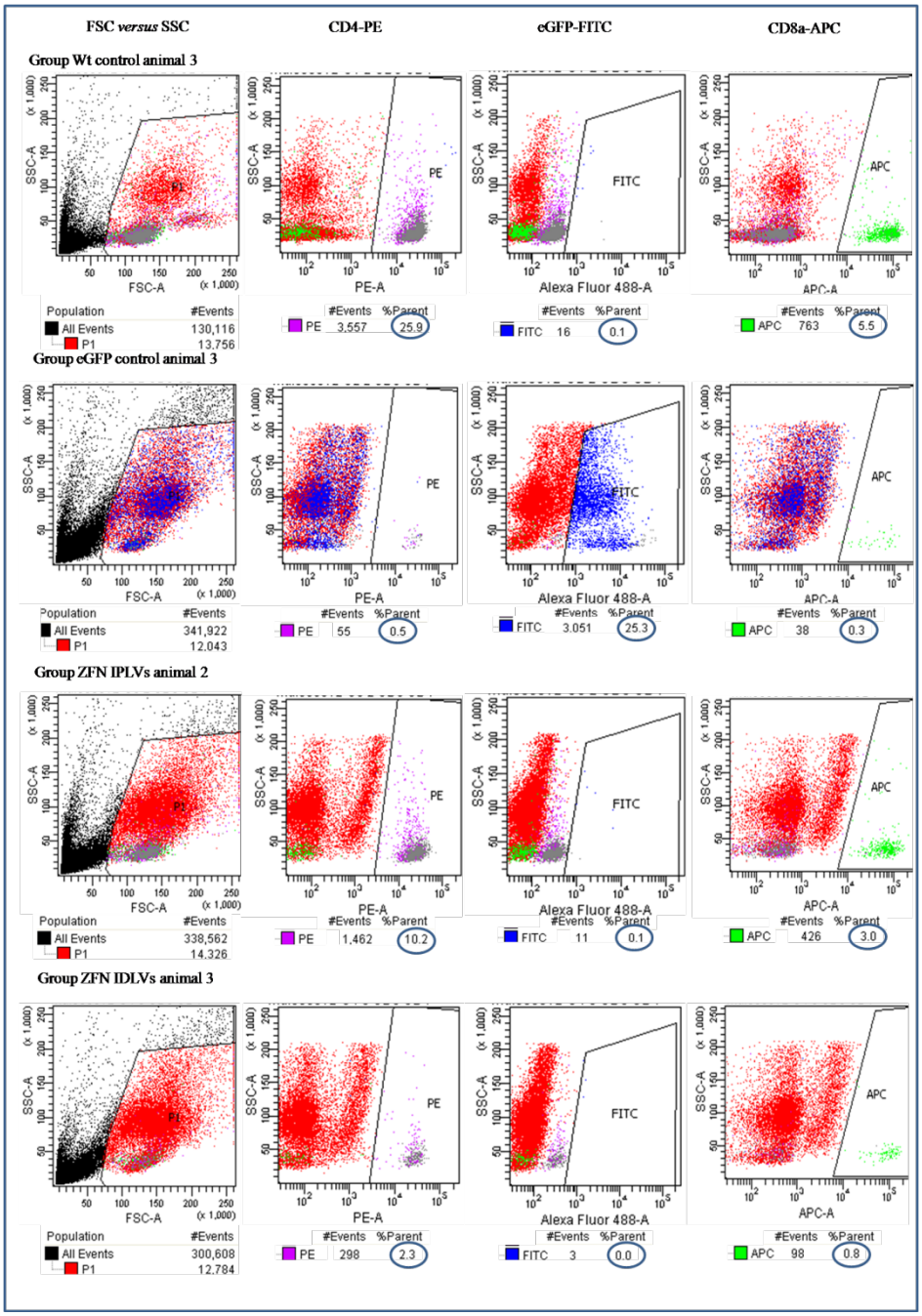



**Figure 6.10: Flow cytometry analysis of CD4 and CD8 populations in peripheral blood of selected samples at 24 weeks post-transplantation.** Blood samples were collected from recipient mice at 24 weeks post-transplantation and multicolour flow cytometric analysis for HSC engraftment was completed for one animal of each group by staining with anti CD4 and CD8a. The animal of group wt control showed ~ 26 % $CD4^+$ and ~ 5.5 % $CD8a^+$, the eGFP control showed ~ 0.5 % $CD4^+$ and ~ 0.3 % $CD8a^+$, the ZFN IPLV responder showed ~ 10 % $CD4^+$ and ~3 % $CD8a^+$ and the ZFN IDLV responder showed ~ 2.3 % $CD4^+$ and ~ 0.8 % $CD8a^+$ T cells.

As a summary, the overall results obtained from staining with CD4 and CD8 at 16 and 24 weeks post-transplantation were plotted in Figure 6.11. The tracking of T-cell reconstitution in peripheral blood of the animal of group wt control showed similar finding of two subpopulations of functional T cell subsets in the two periods matching wt mouse populations, while there was no detected reconstitution of any T cell subpopulations in the eGFP control. On the other hand, in the two responder animals (of groups ZFN IPLV and IDLV) there was a noticeable and gradual increase in the percentages of $CD4^+$ and $CD8^+$ T cell subpopulations in the same periods. In conclusion, T cells reconstitution in the two responder animals was achieved with slow kinetics.

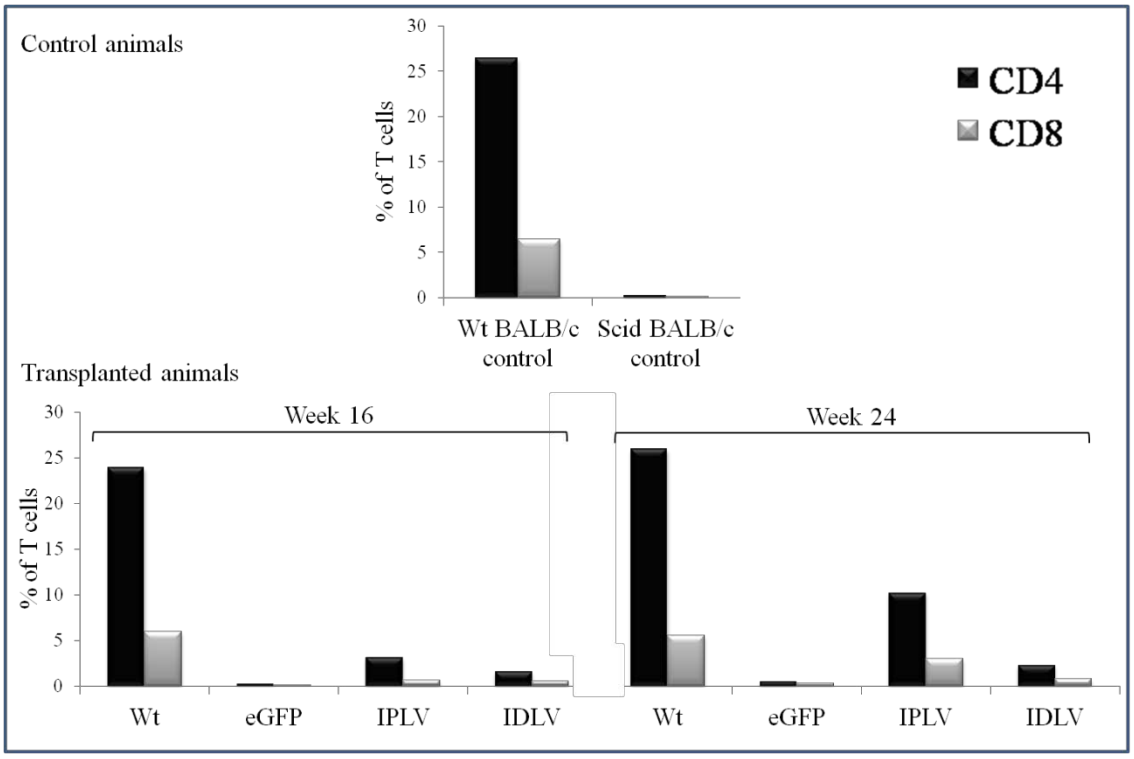



**Figure 6.11: Tracking reconstitution of $CD4^+$ and $CD8^+$ T cell subpopulations at 16 and 24 weeks post-transplantation.** Top panel: control animals of 24 weeks old. Lower panel: transplanted animals. The animal of group wt control showed two subpopulations of mature T cell subsets in the two periods similar to wt mouse control; whereas no T cell reconstitution was observed in the eGFP control animal. The two responder animals (of groups ZFN IPLV and IDLV) clearly demonstrated gradual reconstitution of $CD4^+$ and $CD8^+$ T cell subpopulations.

# Chapter Seven
# Discussion and Conclusions

# 7 Discussion and Conclusions

## 7.1 General discussion

PIDs are rare inherited disorders of the immune system developing from genetic mutations affecting specific immune cells (e.g., T or B lymphocytes, NK cells, neutrophils or antigen presenting cells). PID patients are unable to produce a proper protective immune response, leading to an increased susceptibility to life-threatening infections (Al-Herz *et al*., 2011; Chapel *et al*., 2003). Currently, there are more than 150 different types of PIDs have been identified and characterized which range from mild to severe (Notarangelo *et al*., 2009). More than one million patients around the world are suffering from monogenic PIDs (Modell *et al*., 2011). SCID disorders are human monogenic PID inherited diseases caused by faulty alleles of an individual gene of immune cells leading to the onset of one or more different infections within the first few months of life. These infections are usually serious, and may even be life threatening, they may include pneumonia, meningitis or bloodstream infections. The first case of human DNA-PKcs (*PRKDC*) deficiency has been reported recently representing the first case of human *PRKDC* SCID disorder (van der Burg *et al*., 2009b).

The *PRKDC* SCID patient presented immunologically as classical $T^-$ $B^-$ SCID and was characterized as lacking of V(D)J recombination mechanism which in turn has led to severe decrease in T and B lymphocytes numbers in peripheral blood and a composition of the precursor B-cell compartment in bone marrow. Clinically, the patient was highly vulnerable to opportunistic infections. Since the third month of life, she was suffering from frequent oral candidiasis and lower respiratory tract infections with progressive respiratory distress leading to hypoxia. At the age of 4 months she had developed a large oral aphthous lesion and minimal tonsillary tissue. The patient's fibroblasts were sensitive to ionizing radiation, suggesting that she had a NHEJ defect (van der Burg *et al*., 2009a; van der Burg *et al*., 2009b).

The treatment for PID patients is made available to only some of them by HSCT. Luckily, gene therapy is opening up a new era in medicine, it uses patient's own cells and has the promise to cure many PIDs. The preferred approach applied to introduce precise genetic modification would be made by HR mediated gene repair. This involves harnessing of the cellular repair machinery to induce HR mechanism which mediates

exchange of DNA between a homologous corrective template and the genomic target (Olsen *et al.*, 2010). Thus, HR-mediated gene repair has the potential to regain the original function of the defective gene reducing the risk of insertional mutagenesis and ensuring a sustained treatment because of stable expression level of repaired gene. Hence, it would be preferred to modify the existing faulty gene rather than deliver a new one (Kohn, 2010).

Two possible approaches could be used for gene therapy: the first is *in vivo* therapy to correct faulty-gene cells directly without taking them out from the body; and the second is *ex vivo* therapy to correct faulty-gene cells indirectly by removing them out of the body, purifying, correcting, and then transplanting them back into the body (Connelly *et al.*, 2010). Generally, the major challenge in *ex vivo* gene repair is to improve the natural low HR frequency which makes this strategy too far away to be considered as a therapeutic approach. In this respect, two promising developments have facilitated the task to look for successful possibilities to increase HR rate: the designer ZFNs and efficient gene delivery systems like lentivectors and AAVs.

## 7.2 Thesis hypothesis

In the context of this study, I aimed at developing an *ex vivo* system to improve the efficiency of gene repair in HSCs in order to correct the *Prkdc* mutation of the classical *scid* mouse, a model of human *PRKDC* SCID disorder. The *scid* mouse suffers from a nonsense point mutation in *Prkdc* which causes a truncation of 83 amino acids at C-terminus region of DNA-PKcs (Bosma *et al.*, 1983). To achieve this approach, lentiviral vectors were recruited to deliver ZFNs which create site-specific DSBs at the mouse genome very close to the *scid* point mutation to enhance cellular HR-mediated gene repair machinery. The *ex vivo* treated HSCs were transplanted back into *scid* mice recipients where they are expected to have a selective advantage to proliferate and restore the *scid* mouse immune system.

Although the first *in vivo* gene repair of a mouse model of haemophilia has been done successfully (Li *et al.*, 2011b), there is no animal model of haematopoietic disease that has been corrected yet by gene repair. The results presented in this work highlight the potential of gene repair for future therapy to treat recessive genetic PID disorders. The success of this approach to treat a haematological disorder could be applied to cure

many other inherited diseases that are potentially amenable to correction by gene repair. Nonetheless, a number of challenges had to be addressed before considering a therapeutic approach to rescue the *scid* mouse using combined technology of ZFNs and lentiviral vectors: (i) the ZFN cutting specificity with minimal genotoxic / cytotoxic adverse effects, (ii) lentiviral vectors as a ZFNs delivery system with maximal expression and biosafety, and (iii) selective advantage ability of the corrected cells to repopulate and rescue the mouse immune system.

Previous reports have shown that ZFN technology can efficiently modify human CD34+ HSCs (Holt *et al*., 2010). Moreover, the transplantation techniques nowadays are well-established to manipulate the haematopoietic cells clinically. Thus, the objective of the investigations in this thesis is to explore the use of new ZFNs technology to induce site-specific DSBs at the *scid* mouse *Prkdc* aiming to a significant increase in HR frequency which eventually creates permanent repair. The genes coding for ZFNs and donor template will be delivered by lentiviral vectors to maximise the opportunity for efficient gene expression. The vital aim is to demonstrate proof-of-principle for correction of haematopoietic disease which will contribute to the development of clinically applicable approach to cure patients suffering from PIDs.

## 7.3 ZFN design and delivery platforms

ZFNs are known to introduce sequence-specific DNA DSBs that can be repaired by either NHEJ generating small alterations at targeted genomic sites, or by HR-mediated gene repair in the availability of a homologues DNA sequence. Accordingly, ZFN technology has emerged rapidly and facilitated highly efficient gene disruption in many cell types (Hockemeyer *et al*., 2009; Santiago *et al*., 2008), plant engineering (Tzfira and White, 2005; Zhang *et al*., 2010b) animal models (Doyon *et al*., 2008; Geurts *et al*., 2009b) and even mediated the progress of targeted gene therapy in humans (Holt *et al*., 2010; Perez *et al*., 2008). The state-of-the art ZFN technology could be deigned to target any of the 64 possible DNA 3-base combinations, making ZFN a custom designed molecule for genome surgery (Kohn, 2010). So, theoretically ZFNs could be designed to target any site in the genome including those mutated ones of inherited disease patients. But, practically not all the genomic regions are accessible because of many reasons like compact chromatin structure and DNA modification (Carroll, 2011).

As an example, it has been suggested that cleaving of intact recognition sites of homing endonucleases is prevented by chromatin structure in *S. cerevisiae* during mating-type switching (Rusche *et al*., 2003). The ZFN recognition site in mammalian genomes is at least 18 bp in length which should occur only once, offering a unique site within the genome.

Despite these achievements, a major challenge related to this technology is their imperfect specificity, where unintended DSBs induced at genomic DNA (off-target cutting) may lead to genotoxic and cytotoxic effects (Cornu *et al*., 2008). The binding of ZFN domains to DNA sites could be determined by several techniques like ELISA assays, microarrays and systematic evolution of ligands by exponential enrichment (SELEX) (Zykovich *et al*., 2009). However; it has been reported that ZFN cytotoxicity is mainly ascribed to DNA cleavage rather than domains binding alone (Beumer *et al*., 2006). Importantly, Gabriel *et al*, have shown that ZFN wide spectrum of off-target cutting could be determined by deep sequencing and linear amplification-mediated (LAM)-PCR (Gabriel *et al*., 2011).

ZFNs used in this study were prepared by Sangamo BioScience Inc (California, USA) as collaboration. These nucleases were engineered to encompass 5 and 4 zinc-finger domains in the right and the left ZFN monomers, respectively (Figure 1.8). In principle, at least three fingers in each ZFN monomer are required to offer sufficient affinity for efficient and specific cleavage. More fingers could be added to improve ZFN specificity, as well as affinity, and examples of up to six fingers have been used (Geurts *et al*., 2009a; Hockemeyer *et al*., 2009; Porteus, 2009); however recently it has been confirmed that extra fingers of 5 or 6 in each monomer could lead to significant decrease in the performance of designed ZFNs (Shimizu *et al*., 2011). For maximising the specificity of ZFN cutting and minimising possible off-target cutting, ZFNs designed for this study have included obligatory heterodimeric monomers via mutations in the *Fok* I domain preventing illegal cutting from unwanted homodimers. This strategy has been used earlier by others to improve ZFNs specificity (Miller *et al*., 2007; Szczepek *et al*., 2007). Noteworthy, continuous development in ZFN technology has highlighted that highest ZFN performance can be achieved by providing highest DNA-binding affinity and specificity leading to minimise ZFN off-target related toxicity

(Cornu *et al*., 2008). Therefore, the design of ZFNs used in current study takes into account the maximal ZFN features required both for affinity as well as specificity.

In present thesis, it is expected that the ZFN cutting site could occur anywhere within the nucleotide spacer (of 6 nucleotides) between the two recognition sites of ZFN monomers (see Figure 1.8). The ZFN recognition (binding) and the ZFN cutting sites are located at 20 and 32-38 nucleotides downstream from the *Prkdc scid* point mutation, respectively (see Appendix, Box 1). In the framework of tailoring ZFN recognition site to target mouse *Prkdc*, the point of major importance was to keep the ZFN site as close as possible to *scid* point mutation to ensure that the corrected *scid* site will be within the gene conversion tract, which will be presented during the HR-mediated gene repair. As it is observed earlier that in mammalian cells the gene conversion tracts are generally short; between 100– 200 bp from the DNA break (Elliott *et al*., 1998).

A diagnostic feature of restriction enzyme *Bsa*WI has been introduced via site directed mutagenesis at 2 nucleotides upstream from *scid* point mutation locus in *Prkdc-Bsa*WI template creating only a silent mutation (for exact site see Appendix, Box 2). The advantage of having this site is to be used as a diagnostic tool to certify that possible repaired *Prkdc* had occurred because of incorporation of corrective *Prkdc-Bsa*WI template during HR-mediated gene repair; not by natural spontaneous mutation. Also the presence of this site was used successfully to estimate HR repair frequency in the targeted cells (in *Bsa*WI assay). Moreover, both corrective *Prkdc-Bsa*WI and *Prkdc-neo* templates used in this work were encoded the ZFN recognition site (see Figure 4.8). The idea of incorporation of a mutant ZFN recognition site into the corrective template, to prevent template cutting, was excluded as the main purpose for this study is to correct *scid* point mutation. Adding another mutation, even if it is silent, to the homologous donor sequence could result in reduction of % identity between supplemented template and genomic target, leading to possible significant reduction of HR frequency. Once the ZFN two monomers are dimerised, there will be a continuous cycle of cleaving and rejoining events in their target sites, until a NHEJ repair process generate a mutated ZFN target site (Kandavelou *et al*., 2009).

Different applications have used different approaches to deliver ZFNs into the targeted cells. In plants, ZFNs have been delivered effectively by direct DNA transformation (Cai *et al*., 2009; Shukla *et al*., 2009), agrobacterial transformation under the expression

of a viral promoter (de Pater *et al*., 2009; Zhang *et al*., 2010b), and viral delivery (Marton *et al*., 2010). Direct injection of ZFN mRNA into the embryos of Drosophila (Beumer *et al*., 2008), zebrafish (Foley *et al*., 2009), frog (Young *et al*., 2011) rat (Geurts *et al*., 2009a; Mashimo *et al*., 2010), and mouse (Carbery *et al*., 2010) have been used successfully. Other ZFNs have been introduced into mammalian cells by direct DNA transfection via nucleofection or LipofectAMINE (Orlando *et al*., 2010; Perez *et al*., 2008; Urnov *et al*., 2005), and viral delivery (Li *et al*., 2011b; Lombardo *et al*., 2007). Interestingly, it has been demonstrated very recently the possibility of direct delivery of purified ZFN proteins into targeted cells (Gaj *et al*., 2012).

In current study, the ZFNs are delivered by IDLV platforms while the IPLVs were used alongside for comparison and optimisation. A one useful approach to bypass inefficient transfection problems (Ellis *et al*., 2012), ensure transient gene expression, and enhance HR therapeutic levels; is to use lentiviral vectors as a delivery strategy. It has been shown earlier that lentiviral vectors are able to transduce quiescent human HSCs (Case *et al*., 1999). Thus, using of lentiviral vectors would be beneficial because: firstly, lentiviral vectors are able to efficiently transduce quiescent cells, like HSCs, as they incorporate NLS which facilitate entry of PIC into nuclei of non dividing cells. Secondly, the episomal nature of IDLVs to generate transcriptional functional elements rather than integrate into the genome made them safer compared to other lentiviral vectors. This is be of particular importance for the purpose of delivering ZFNs to ensure transient expression leading ultimately to reduce ZFNs overexpression and minimise possible cytotoxic effect because of off-target cutting. And thirdly, the lentiviral vectors used in present study to deliver ZFNs are self-inactivating third-generation human immunodeficiency virus-1 (HIV-1)-based vectors have been known with their improved biosafety (Dull *et al*., 1998; Yanez-Munoz *et al*., 2006; Zufferey *et al*., 1998).

Three research groups have demonstrated that recruiting of IDLVs to induce site-specific DSBs can be effective method to achieve high HR-mediated gene repair frequency in a range of human and mouse cell types (Cornu and Cathomen, 2007; Lombardo *et al*., 2007; Okada *et al*., 2009). Additionally, combination of IDLVs and ZFNs could result in approaching the threshold that would be therapeutically useful for PIDs (Lombardo *et al*., 2007). Hence, the possibility of therapeutic use of ZFN-mediated gene repair in mammalian cells has become more hopeful and it seems that

some monogenic PIDs could be cured by application of this strategy. Such progress has led to anticipate that the first monogenic inherited disease to be cured would be SCID disorders (Hatada *et al*., 2000; Miller *et al*., 2007).

The success of ZFN recruitment depends on their abilities to induce site-specific DSBs at their targeted loci. Both repair mechanisms of HR and NHEJ are presented in essentially all cells and they are likely to be presented in most circumstances. Nevertheless; which repair mechanism will be used in DSBs repair depends on many factors like (i) cell cycle status (Branzei and Foiani, 2008), (ii) availability of DNA repair proteins (Cann and Hicks, 2007), and (iii) availability of homologous donor templates (Shrivastav *et al*., 2008). The alterations that could influence some of HR or NHEJ repair components, could lead to swap from one to another pathway. As an example, it is known that DNA ligase plays a vital role in the NHEJ repair pathway, but in its absence the repair pathway will be diverted mainly toward HR (Carroll, 2011).

The *scid* HSC progenitors are the main target cells to be corrected in present study; however for convenience and practicality the BALB/c 3T3 *scid* fibroblasts are used to optimise experiments of demonstration of ZFN cutting, gene targeting and gene repair. I started the work in this project with a problem of extensive cellular senescence of *scid* fibroblasts. These cells had showed limited life-span and they were dying upon cloning before completion of the intended experiment. Therefore, and in order to overcome this problem they were immortalised by transduction with retroviral vectors encoding m*Tert* and *Pac*. Telomerase is an RNA-dependent DNA polymerase that catalyzes the addition of telomeric repeat sequences to chromosome ends, and telomeres shorten with successive cell divisions in the absence of telomerase (Harrington *et al*., 1997). Importantly, the temporal start of senescence is related firmly to telomere length (Harley *et al*., 1990). Cellular senescence has been bypassed by immortalisation with retroviral vectors expressing human TERT in human cells (Vaziri and Benchimol, 1998) or *Tert* in mouse cells (Kawai *et al*., 1994). Upon their immortalisation, I noticed that m*Tert scid* fibroblasts possessed considerably higher clonal efficiency and that such clonal populations were able to grow and expand to complete the required experiments.

Previous work in current study has begun prior to my arrival by designing of two long corrective donor templates both coding for wild type homologous sequences, but they differ in their additional markers. The first one carried the *neo* cassette to be used as a

positive selective marker in *in vitro* system optimisation; this template is referred to as *Prkdc-neo* template. The second encoded a novel site for the enzyme *Bsa*WI (introduced via site-directed mutagenesis causing only silent mutation) to be used as a diagnostic feature of occurrence of gene repair. Subsequently, the last template was replaced with a shorter template coding for the same *Bsa*WI site and called *Prkdc-Bsa*WI template, which already used for gene repair experiments *in vivo* and *ex vivo.* The short template was used because of two reasons: the first, it was easier to clone it together with ZFN monomers into the same lentiviral backbones, and the second, it was more practical to detect gene repair events in targeted cells by PCR amplification using external primers to the shorter template (*Bsa*WI assay).

The long corrective template with *Bsa*WI site (that was replaced afterwards) was incorporated into lentiviral backbones to construct pCCLsc*Prkdc*Hind*Bsa*WIF and pCCLsc*Prkdc*Hind*Bsa*WIR. Both constructs were incorporating the same sequence (*Hin*dIII *Bsa*WI fragment), which is coding for the wild type sequence around the *scid* point mutation in the *Prkdc* gene and *Bsa*WI site. However, there was only one difference between them which is the orientation of inserted Hind *Bsa*WI fragment; one construct carried the insert in forward orientation while the other in reverse. These two constructs prepared to be used as donor templates in HR-mediated gene repair; but eventually I did not use them anywhere as I replaced them by *Prkdc-Bsa*WI template. On the other hand, three different strategies (data not shown) were carried out in order to clone the *Prkdc-neo* template into lentiviral backbones to construct pCCLsc*Prkdc*Hind*Pml*I*neo*; but I was not able to get it. This could be attributed to the large size of insert and inconsistency between plasmid electronic file and real plasmid as sometimes I ended up with fragments which are not expected by vector NTI software. The cloning of obligatory heterodimeric ZFNs into lentiviral backbones with and without *Prkdc-Bsa*WI template has been carried out successfully.

I have produced ZFNs, ZFNs/*Prkdc-Bsa*WI template and CMV-eGFP third-generation HIV-1 based lentiviral vectors, and SFFV-eGFP second-generation HIV-1-based lentiviral vectors by transient transfection of 4 or 3 plasmids, respectively, into HEK-293T cells using the calcium phosphate coprecipitation method (as explained in section 2.2.7). HEK-293T cells were used as producer cells because of their high transfect-ability to calcium phosphate. All prepared lentiviral vectors were titrated in HeLa cells

by real-time quantitative PCR, which is most reliable method to titrate HIV-1 based lentiviral vectors as it measure all HIV-1 based lentiviral vectors cDNA varieties (Yoder and Fishel, 2008). Those eGFP prepared vectors were additionally titrated by flow cytometry analysis.

As all ZFNs used in current study were FLAG-tagged and in order to detect of ZFN expression in *scid* m*Tert* fibroblasts from original pVAX plasmids, lentiviral plasmids, and produced lentiviral vectors, Western blotting assay was used. A successful ZFN monomer detection was observed from both second and third version ZFN pVAX constructs. Also, ZFN expression was detected from all cloned monomeric ZFN lentiviral (pRRL and pCCL) backbones. However, the expression from dimeric ZFN constructs was not detectable because of the lack of a N2A self-processing polyprotein signal which helps to separate ZFN monomers during transcription. In fact, first ZFN version had included this signal, but next versions (second and third) did not includ it. Thus, ZFN dimeric constructs were not used in any next experiments. On the other hand, I have detected ZFN expression from all produced lentiviral vectors and I also have not noticed any difference between ZFN expression under the control of CMV and SFFV promoters. However; ZFN expression from IPLVs was stronger than that from IDLVs, which can be explained if expression from IPLVs is more efficient than that of IDLVs in these highly proliferating cells (Wanisch and Yanez-Munoz, 2009). As an attempt to detect ZFN expression in lin$^-$ *scid* HSCs, the transduction with ZFN lentiviral vectors was conducted and followed by detection with internal antigen labelling (data not shown). This assay is commonly used to check the expression of surface and intracellular antigens. Unfortunately, the flow cytometric analysis showed that there were no detectable signals above the background which could be attributed to weak ZFN signals.

Three versions of *Prkdc* ZFNs were received from Sangamo; the first contained a first version of the ZF binding domains and non-optimised wild-type *Fok* I domains which allow homodimerisation; the second had optimised ZF domains and wild-type *Fok* I domains; and the third version contained the optimised ZF domains and mutated, obligatory heterodimeric *Fok* I domains. Optimisation of ZF domains by Sangamo was on the basis of their binding activity to the cognate sequence. A gene targeting assay was conducted in order to select the best *Prkdc* ZFN version that could induce higher

levels of DSB-mediated HR which can be detected by number of G-418r CFUs using the three versions of *Prkdc* ZFNs and *Prkdc-neo* template (section 4.4). As it was expected, the results clearly demonstrated that the combination of optimised ZF domains with obligatory heterodimeric monomers (third version ZFNs) produced the highest number of G-418r CFUs (10-fold increase) suggesting highest DSB generation. This finding has agreed with previous report when similar system was used (Olsen *et al*., 2010). Different approach using same strategy (co-transfection of ZFN monomers and donor plasmid) has showed potentially relevant frequencies of ZFN-mediated gene targeting in a variety of primary cells of mouse model of a generic recessive genetic disease (Connelly *et al*., 2010).

Ideally gene targeting/repair events could to be verified easily by Southern blotting assay, as this assay can detect the presence of a known DNA fragment within a sample, as well as to inspect the integration status of the sequence, by probing DNA digested with a restriction enzyme. In this context, two different radioactive probes were designed (data not shown) and then used to attempt detection of gene targeting at G-418r CFUs using Southern blotting. Unfortunately, after few trials with each probe I could not observe any signal for gene targeting; a matter that could be attributed to that signals from those probes were very poor. Because of that failure, I conducted other alternative experiments to demonstrate gene targeting events at targeted locus in G-418r fibroblasts by PCR based analysis. By designing targeting construct internal (forward) and external (reverse) primers (Figure 4.10), I was able to detect an amplification of a 1335-bp product in 20 out of 36 examined samples. This is represented a clear evidence for ZFN-mediated gene targeting at the mutated *Prkdc* locus. For further confirmation, I carried out an enzymatic digestion using successful amplified samples to double check whether the right amplicons had been amplified. All 20 amplicons subjected to digestion with *Bgl*II showed the presence of fragments of 746 and 589bp suggesting that right amplification. Additionally I did one more confirmation step; randomly selected sample was subjected to further enzymatic digestion with *Hin*dIII and *Xho*I generating right expected fragments. The finding of that 20 out of 36 G418r clones were targeted although the increase in their frequency was 10-fold in previous experiment (above paragraph) could be explained by that the 36 G418r clones were selected randomly out from two separated experiments (18 clones each) allowing the collection of some non-targeted G418r clones and eventually leading to reduce the targeting frequency.

In the light of above gene targeting results, only the third version *Prkdc* ZFNs (optimised obligatory heterodimeric) was cloned into lentiviral-backbones and eventually used to produce IPLVs and IDLVs. Additionally lentiviral-backbones carrying both a ZFN monomer and the *Prkdc-Bsa*WI template were generated and used to produce IDLVs. Therefore; the first and second ZFN versions were not used anywhere in next experiments.

## 7.4 ZFN cutting specificity

I was particularly interested in further characterising of ZFN abilities to generate site-specific cuttings at their genomic target site. ZFN-induced DSBs would be repaired usually by NHEJ in absence of homologous donor template leading to the introduction of InDels, which can be detected by the Surveyor *Cel*-I nuclease assay. In *scid* m*Tert* fibroblasts, my results revealed that there was an effective dose- and dimer-dependent ZFN-mediated cutting at the target site. Also, correlating with the increased expression level from IPLVs noticed in the Western blot, InDel frequencies from IPLVs were higher than those from IDLVs for both CMV or SFFV-driven ZFNs, such finding would be understandable in these proliferative cells. ZFN expression under CMV or SFFV promoters reached similar levels. In lin$^-$ *scid* HSCs, I was only able to observe effective dose- and dimer-dependent ZFN-mediated cutting at the target site with ZFN IPLVs driven by SFFV promoter, suggesting that in this experiment only this configuration and promoter mediated sufficient target cutting.

Because of limited sensitivity of Surveyor *Cel*-I nuclease assay (I showed that its minimal detectable modification was 5% in section 4.5.1), I was unable to detect ZFN site-specific cuttings in ZFN IDLVs controlled by SFFV promoter in lin$^-$ *scid* HSCs. Nonetheless, I detected ZFN site-specific cuttings by deep-sequencing technology instead. Deep-sequencing results showed that there are effective dose- dependent ZFN-mediated cutting when SFFV-driven ZFN IDLVs used. Although detected ZFN site-specific cuttings from IDLV SFFV was not as high as (maximum ~ 0.5 % at high dose) that observed in IPLVs, yet still informative to confirm that *Prkdc* ZFNs are able to induce efficient site-specific cuttings at their targeted loci. Noteworthy, the results of using IDLVs with SFFV-driven eGFP in HSCs confirmed that expression levels were

extremely poor compared to the integrating configuration, confirming the unsuitability of this promoter in IDLVs

Using the Surveyor *Cel*-I nuclease assay I have also been unable to detect ZFN site-specific cutting in HSCs using IPLVs or IDLVs driven by the CMV promoter. These results could be explained by the fact of CMV is not compatible for robust expression in hematopoietic lineages in contrast to its very good compatibility in other cell lines like HEK293 and HeLa (Logan *et al*., 2002). It is worth mentioning that the SFFV promoter has been broadly used to regulate transgene expression within retroviral vectors, especially due to its high activity in hematopoietic cells (Demaison *et al*., 2002; Kung *et al*., 2000; Zhang *et al*., 2010a).

Interestingly, one of major questions that arose in current work was: how these DNA-PKcs deficient cells can repair their induced-DSBs by NHEJ mechanism? (as these NHEJ-repaired DSBs were observed by Surveyor *Cel*-I nuclease assay). This repair mechanism could be attributed to the assumption of that in the case of absence of one or more of NHEJ required factors (like Ku DNA, DNA-PKcs, DNA ligase IV XRCC4, and a number of DNA end-processing factors), there is an ‘‘alternative’’ NHEJ pathway could take place instead allowing the cells to repair their DSBs even in the absence of classical NHEJ (Brunet *et al*., 2009; Carroll, 2011; Lieber *et al*., 2003).

## 7.5 ZFN-mediated gene repair

I expanded my research to detect HR events (by *Bsa*WI assay) in targeted *scid* fibroblasts and HSCs using a silent diagnostic restriction site for the enzyme *Bsa*WI, previously introduced via site-directed mutagenesis two nucleotides upstream from the *scid* mutation site in the *Prkdc*-*Bsa*WI template. The successful gene targeting attempt could be followed by the detection for cells carrying the new DNA using *Bsa*WI diagnostic feature. In *scid* m*Tert* fibroblasts, the results demonstrated that gene repair could be detected in SFFV-driven ZFNs in both IPLVs (maximum 7.7 % at MOI 1000) and IDLVs (maximum 4.7 % at MOI 5000), and was vector dose-dependent. Using the same assay, no gene repair was observed in the case of CMV-driven ZFNs used in either IPLVs or IDLVs. While in lin$^-$ *scid* HSCs, the results showed that gene repair could only be detected in SFFV-driven ZFNs in IPLVs (8.3 % at MOI 100). In fact, the latter gene repair calculations could be an overestimated result, as the diagnostic band of

1068 bp was about to be seen. It is important to state that in both *scid* fibroblasts and HSCs, the calculations of gene repair results were based on a formula with a correction factor to obtain more accuracy as the smallest band of 640 bp was hard to be detected because the overlapping smear on the gel.

As deep-sequencing results suggested earlier that ZFN site-specific cuttings would be low generally in the case of SFFV-driven ZFNs used in IDLVs, the undetectable gene repair in this case might be justifiable. Lack of detection of gene repair in CMV-driven ZFNs would not be unexpected because of that CMV promoter is not as good as SFFV promoter to control lentiviral vector gene expression in hematopoietic cells. However, for the *ex vivo* gene repair and transplantation experiment I used CMV-driven ZFN IDLVs because this configuration has showed the higher eGFP expression levels among all vectors and promoters that were tested in HSCs.

To this end, from above results of Surveyor *Cel*-I nuclease assay to detect ZFN site-specific cuttings and *Bsa*WI assay to detect HR events in targeted locus, it can be concluded that our third *Prkdc* ZFN version can cut the target locus and induce gene repair to correct *scid Prkdc*.

## 7.6 Restoration of functional *Prkdc*

The next important question to be addressed was whether correction of *Prkdc* gene at DNA level had led to restore gene function and produce functional DNA-PK? To answer this, two experiments were conducted using gene-repaired *scid* m*Tert* fibroblasts. The first one is concentrated on detection of DNA-PK activity using SignaTECT DNA-dependent protein kinase assay. The results demonstrated that there is a vector dose-dependent restoration of DNA-PK activity in the case of SFFV-driven ZFN IPLVs or IDLVs were used, matching the results of the BsaWI assay. The other experiment was carried out to explore the functionality of enzyme DNA-PK through the response to DNA damaging agents like melphalan. Melphalan can induce DSB which will be repaired by NHEJ and importantly DNA-PK is a determinant in the cellular response to this DNA damaging agents (Muller *et al*., 2000). The results revealed that the number of enriched melphalan-resistant CFUs was remarkably vector dose-dependent. Furthermore, these melphalan-resistant CFUs were subjected to *Bsa*WI assay to investigate whether melphalan-mediated enrichment has correlated with gene

repair at the DNA level. The results showed that there is (1.5-3) fold increase in gene repair frequency detected after enrichment with melphalan, compared to that prior to drug treatment. Alternatively, what I could have done is to isolate individual melphalan-resistant clones to try and identify one that has undergone correction and then show *scid* correction clearly by Southern blotting. Generally, these results have suggested that gene-repaired cells had a selective advantage over uncorrected cells in the presence of melphalan. Altogether, the results obtained from above two experiments clearly confirmed that correction of *Prkdc* has led to restoration of functional DNA-PK.

## 7.7 *Ex vivo* gene repair and transplantation

HSC progenitors are stem cells with self-renewal and multipotency abilities; their multilineage differentiation generates all hematopoietic cells of the myeloid and lymphoid lineages including erythrocytes, platelets, macrophages, T-cells and B-cells. They emerge early during embryogenesis and maintain hematopoiesis throughout the whole lifetime of the organism (Medvinsky *et al*., 2011). HSC progenitors are very rare cells, for example they can be found in adult human bone marrow at a frequency of only one in thousands of cells. While from one mouse, only about a thousand HSC progenitors can usually be obtained. Importantly, HSCs are distinguished from other cells in myeloid tissue by their lacking of the surface markers found in more mature blood cells ($lin^-$) but do express CD117 (c-kit) and stem cell antigen 1 (Sca-1). Their isolation from the surrounding cell populations is therefore based on fluorescent or magnetic labelling and following removal of lineage-positive cells. However and in spite of advanced isolation techniques, the HSCs separated are still heterogeneous population. It has been reported that after murine transplantation, only 30 % - 50 % of potential HSCs were found to be real HSCs after single cell molecular analysis (Osawa *et al*., 1996; Schroeder, 2010). Furthermore in a landmark paper, it has been interestingly shown that a after transplantation of single HSCs into mice, > 25% of stem cells are capable of repopulating the hematopoietic system (Dykstra *et al*., 2007).

Because HSCT procedure has became an established treatment for different patients suffering from various haematological disorders, the frequency of HSCT has increased worldwide (Mohty and Apperley, 2010). The source of HSCs used for HSCTs could be bone marrow, cord blood, or mobilized peripheral blood (Zibara *et al*., 2012). In present

study, bone marrow lin$^-$ HSCs were routinely purified by magnetic cell depletion and grown without passaging in serum free comprehensive medium containing a cocktail of cytokines for *in vitro* culture. Cytokines human IL-6 and murine Flt3 have been added to promote stem cell proliferation in culture, and the murine SCF was added to promote survival rather than cell division (Broudy, 1997).

As *scid* lin$^-$ HSCs were the main target cells to be corrected here in current study, I have particularly concerned to achieve the highest possible transduction efficiency in these cells. For that, I have assessed different vectors, configurations and promoters. I carried out lin$^-$ HSCs transduction with SFFV-driven eGFP IPLVs and IDLVs, CMV-driven eGFP IPLVs and IDLVs, control (no external) promoter-driven eGFP IDLVs, GFAP-driven eGFP IDLVs, PGK-driven eGFP IDLVs, SV40-driven eGFP IDLVs, ASFV P11.5-driven eGFP IDLVs, ASFV P54-driven eGFP IDLVs and three derivatives of Ad5 CMV-driven eGFP adenoviral vectors. The overall results from these experiments suggested that SFFV IPLVs and CMV IDLVs were most efficient; therefore I used these optimised conditions for *ex vivo* transduction.

In the main experiment of present study "rescue of *Prkdc* deficiency by *ex vivo* gene repair and transplantation", the recipient mice were divided into 4 groups of five animals each: Wt and eGFP control groups, and ZFN IPLVs and ZFN IDLVs experimental groups. The purified *scid* lin$^-$ HSCs in eGFP control group were transduced with lentiviral vectors encoding the eGFP. For ZFN IPLVs and ZFN IDLVs groups, the purified *scid* lin$^-$ HSCs were transduced with lentiviral vectors encoding the *Prkdc-Bsa*WI template and ZFNs. The HSCs transduction was about 16 h then followed by treated cells injection into sublethally irradiated recipients. The transduction efficiency in eGFP control group 72 h post-transduction was ~ 38 % at MOI 400. Following *ex vivo* gene repair and upon intravenous injection in transplantation experiment, the corrected cells will be circulated and finally localised at specific body compartments. The injected cells are most likely to be engrafted in bone marrow cavity and spleen, which possess the accommodating stromal elements essential for effective proliferation and hematopoiesis (Samlowski and Daynes, 1985).

To address the main question that arose initially: do the gene-repaired cells have the ability to rescue the mouse immune system? Firstly, the vast majority of mutated native HSCs will be eliminated by sub-lethal irradiation facilitating the expansion and

proliferation of corrected cells with their selective advantage. This ambitious aim theoretically needs just one corrected HSC to repopulate the entire mouse immune system; however the applicability of this fact is quite challenging, taking into consideration many limiting factors including: the original rarity of HSCs in the *scid* mice, the purity of $lin^-$ *scid* HSCs after depletion procedure, the transduction efficiency of corrective template and ZFN lentiviral vectors, all requirements for HR-mediated gene repair mechanism, the efficiency of transplantation and injection procedure, and corrected HSCs seeding efficiency. Secondly, as this project based on gene repair principle meaning that only low percentage of manipulated cell population will be corrected. Subsequently, this has led to face a problem of injection of maximal cell number into recipients in order to maximise the opportunity of higher numbers of transplanted cells. For that, 1 X $10^6$ cells were injected back into recipients. In comparison with previous study to correct other mouse *scid* model, this cell number is considered higher as other study based on gene addition strategy (Huston *et al*., 2011).

In order to track the reconstitution of mouse immune system, all animals were bled at 9, 16, and 24 weeks post-transplantation and PBMCs were stained with different antibodies directed against lymphoid (T and B) and myeloid (monocyte/macrophage, NK and granulocytes) cell CD surface receptors. The observed gradual increase in T cell populations throughout 24 weeks after transplantation in two responder animals (one from each group of ZFN IPLV and IPLV) is considered as a dramatic finding showing that possible immune system restoration has begun, albeit possibly at different speeds. On the other hand, B cells restoration during same period is not observed which can be attributed to that B cells restoration is slower and requires higher numbers of transplanted cells, at least in the *Il2rg* model (Huston et al., 2011). The overall conclusion that could be made here is that transplanted ZFN-mediated gene-repaired cells could lead to slow rescue of the mouse immune system. In this regard, an additional experiment could be carried out by probing for Y-chromosome-specific markers in non-lymphoid lineages in order to assay for a positive control for engraftment. This experiment would be helpful to track the origin (donors or recipients) of all non-lymphoid cells in responder animals.

Altogether, the work presented here showed the feasibility of recruiting ZFN technology to mediate gene repair to correct classical *scid* mouse point mutation.

Current work has provided a proof of principle to restore both original *Prkdc* genotype and phenotype. Further experiments would be required to ascertain of both *Prkdc* correction on DNA level and kinase function after *ex vivo* gene repair transplantation. As a closing remark, all the obtained results suggested that *scid* mouse model is possibly rescued representing the first animal model of haematopoietic disease to be corrected by gene repair. The highly importance of such achievement could contribute to the ongoing efforts to treat monogenic inherited diseases.

## 7.8 Conclusions

The experiments in this Thesis have revealed the following important conclusions:

- *Scid* fibroblasts transduced with *mTert-pac* retroviral vectors are resistant to puromycin, and can be cloned and expanded, suggesting immortalisation.
- ZFN monomers are expressed efficiently after transfection or transduction of *scid* fibroblasts, and less so in HSCs.
- ZFN-mediated gene targeting with a *neo*-containing construct results in a significant increase in G418r CFUs, suggestive of DSB-mediated HR, with a ZF-optimised, obligatory heterodimer ZFN. Clonal analysis of G418r CFUs by PCR indicated that ~ 60 % of the randomly selected samples have undergone gene targeting.
- ZFN site-specific cutting and ZFN-mediated gene repair at the *Prkdc* locus has been demonstrated with IPLVs and IDLVs. In HSCs only the former configuration produces detectable gene repair by gel assay.
- After *Prkdc* correction, DNA-PK activity is restored and corrected fibroblasts show enhanced resistance to the DNA damaging drug melphalan.
- Preliminary results of the *ex vivo* gene repair and transplantation experiment indicate potential rescue of the T-cell compartment in a fraction of the *scid* transplant recipients.

## 7.9 Future work

- Second transplantation experiment using animals of first transplantation (one animal of each control groups of wt and eGFP, and two responder animals) has been conducted already. It is hoped that the corrected HSC progenitors from the first transplantation will be able to rescue recipients in the second transplantation. Such rescue would be considered a further demonstration of the presense of gene-repaired HSCs. The results are still in processing.

- Molecular analysis of gene repair events at the DNA level in cultured *scid* fibroblasts and HSCs, and in PBMCs of responder animals, by deep sequencing. This analysis is required to provide a clear evidence of the correction of the *scid* point mutation which will be a formal demonstration of *scid* genotype repair. The samples were prepared and shipped to Dr. Manfred Schmidt / Germany for analysis.

- Analysis of restoration of DNA-PK activity in PBMCs of responder animals in order to confirm the correlation between *scid* genotype repair and phenotype restoration.

- Analysis of off-target ZFN cutting in cultured *scid* fibroblasts and HSCs, and in PBMCs of responder animals, using PCR amplification and deep sequencing of the 10 most likely off-target sites selected by SELEX protocol. The samples were prepared and shipped to Sangamo for analysis.

# 8 Bibliography


Aaronson, S.A., Todaro, G.J., 1968. Development of 3T3-like lines from Balb-c mouse embryo cultures: transformation susceptibility to SV40. J Cell Physiol 72, 141-148.

Aguila, J.R., Liao, W., Yang, J., Avila, C., Hagag, N., Senzel, L., Ma, Y., 2011. SALL4 is a robust stimulator for the expansion of hematopoietic stem cells. Blood 118, 576-585.

Aiuti, A., Bachoud-Levi, A.C., Blesch, A., Brenner, M.K., Cattaneo, F., Chiocca, E.A., Gao, G., High, K.A., Leen, A.M., Lemoine, N.R., McNeish, I.A., Meneguzzi, G., Peschanski, M., Roncarolo, M.G., Strayer, D.S., Tuszynski, M.H., Waxman, D.J., Wilson, J.M., 2007. Progress and prospects: gene therapy clinical trials (part 2). Gene Ther 14, 1555-1563.

Aiuti, A., Roncarolo, M.G., 2009. Ten years of gene therapy for primary immune deficiencies. Hematology Am Soc Hematol Educ Program, 682-689.

Al-Herz, W., Bousfiha, A., Casanova, J.L., Chapel, H., Conley, M.E., Cunningham-Rundles, C., Etzioni, A., Fischer, A., Franco, J.L., Geha, R.S., Hammarstrom, L., Nonoyama, S., Notarangelo, L.D., Ochs, H.D., Puck, J.M., Roifman, C.M., Seger, R., Tang, M.L., 2011. Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency. Front Immunol 2, 54.

Al-Wahiby, S., Wong, H.P., Slijepcevic, P., 2005. Shortened telomeres in murine scid cells expressing mutant hRAD54 coincide with reduction in recombination at telomeres. Mutat Res 578, 134-142.

Alemany, R., Curiel, D.T., 2001. CAR-binding ablation does not change biodistribution and toxicity of adenoviral vectors. Gene Ther 8, 1347-1353.

Alexander, B.L., Ali, R.R., Alton, E.W., Bainbridge, J.W., Braun, S., Cheng, S.H., Flotte, T.R., Gaspar, H.B., Grez, M., Griesenbach, U., Kaplitt, M.G., Ott, M.G., Seger, R., Simons, M., Thrasher, A.J., Thrasher, A.Z., Yla-Herttuala, S., 2007. Progress and prospects: gene therapy clinical trials (part 1). Gene Ther 14, 1439-1447.

Arhel, N.J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prevost, M.C., Allen, T.D., Charneau, P., 2007. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. EMBO J 26, 3025-3037.

Arnould, S., Perez, C., Cabaniols, J.P., Smith, J., Gouble, A., Grizot, S., Epinat, J.C., Duclert, A., Duchateau, P., Paques, F., 2007. Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. J Mol Biol 371, 49-65.

Azzouz, M., Le, T., Ralph, G.S., Walmsley, L., Monani, U.R., Lee, D.C., Wilkes, F., Mitrophanous, K.A., Kingsman, S.M., Burghes, A.H., Mazarakis, N.D., 2004. Lentivector-mediated SMN replacement in a mouse model of spinal muscular atrophy. J Clin Invest 114, 1726-1731.

Banasik, M.B., McCray, P.B., Jr., 2009. Integrase-defective lentiviral vectors: progress and applications. Gene Ther 17, 150-157.

Bayer, M., Kantor, B., Cockrell, A., Ma, H., Zeithaml, B., Li, X., McCown, T., Kafri, T., 2008. A large U3 deletion causes increased in vivo expression from a nonintegrating lentiviral vector. Mol Ther 16, 1968-1976.

Bayo-Puxan, N., Cascallo, M., Gros, A., Huch, M., Fillat, C., Alemany, R., 2006. Role of the putative heparan sulfate glycosaminoglycan-binding site of the adenovirus type 5 fiber shaft on liver detargeting and knob-mediated retargeting. J Gen Virol 87, 2487-2495.

Bayo-Puxan, N., Gimenez-Alejandre, M., Lavilla-Alonso, S., Gros, A., Cascallo, M., Hemminki, A., Alemany, R., 2009. Replacement of adenovirus type 5 fiber shaft heparan sulfate proteoglycan-binding domain with RGD for improved tumor infectivity and targeting. Hum Gene Ther 20, 1214-1221.

Beamish, H.J., Jessberger, R., Riballo, E., Priestley, A., Blunt, T., Kysela, B., Jeggo, P.A., 2000. The C-terminal conserved domain of DNA-PKcs, missing in the SCID mouse, is required for kinase activity. Nucleic Acids Res 28, 1506-1513.

Belfort, M., Roberts, R.J., 1997. Homing endonucleases: keeping the house in order. Nucleic Acids Res 25, 3379-3388.

Benveniste, P., Cantin, C., Hyam, D., Iscove, N.N., 2003. Hematopoietic stem cells engraft in mice with absolute efficiency. Nat Immunol 4, 708-713.

Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J.K., Carroll, D., 2006. Efficient gene targeting in Drosophila with zinc-finger nucleases. Genetics 172, 2391-2403.

Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., Carroll, D., 2008. Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. Proc Natl Acad Sci U S A 105, 19821-19826.

Blunt, T., Finnie, N.J., Taccioli, G.E., Smith, G.C., Demengeot, J., Gottlieb, T.M., Mizuta, R., Varghese, A.J., Alt, F.W., Jeggo, P.A., et al., 1995. Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine scid mutation. Cell 80, 813-823.

Boch, J., 2011. TALEs of genome targeting. Nat Biotechnol 29, 135-136.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., Bonas, U., 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326, 1509-1512.

Bosma, G.C., Custer, R.P., Bosma, M.J., 1983. A severe combined immunodeficiency mutation in the mouse. Nature 301, 527-530.

BouHamdan, M., Duan, L.X., Pomerantz, R.J., Strayer, D.S., 1999. Inhibition of HIV-1 by an anti-integrase single-chain variable fragment (SFv): delivery by SV40 provides durable protection against HIV-1 and does not require selection. Gene Ther 6, 660-666.

Branzei, D., Foiani, M., 2008. Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol 9, 297-308.

Broudy, V.C., 1997. Stem cell factor and hematopoiesis. Blood 90, 1345-1364.

Brunet, E., Simsek, D., Tomishima, M., DeKelver, R., Choi, V.M., Gregory, P., Urnov, F., Weinstock, D.M., Jasin, M., 2009. Chromosomal translocations induced at specified loci in human stem cells. Proc Natl Acad Sci U S A 106, 10620-10625.

Buchschacher, G.L., Jr., Wong-Staal, F., 2000. Development of lentiviral vectors for gene therapy for human diseases. Blood 95, 2499-2504.

Burns, J.C., Friedmann, T., Driever, W., Burrascano, M., Yee, J.K., 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. Proc Natl Acad Sci U S A 90, 8033-8037.

Butler, S.L., Hansen, M.S., Bushman, F.D., 2001. A quantitative assay for HIV DNA integration in vivo. Nat Med 7, 631-634.

Butler, S.L., Johnson, E.P., Bushman, F.D., 2002. Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. J Virol 76, 3739-3747.

Cai, C.Q., Doyon, Y., Ainley, W.M., Miller, J.C., Dekelver, R.C., Moehle, E.A., Rock, J.M., Lee, Y.L., Garrison, R., Schulenberg, L., Blue, R., Worden, A., Baker, L., Faraji, F., Zhang, L., Holmes, M.C., Rebar, E.J., Collingwood, T.N., Rubin-Wilson, B., Gregory, P.D., Urnov, F.D., Petolino, J.F., 2009. Targeted transgene integration in plant cells using designed zinc finger nucleases. Plant Mol Biol 69, 699-709.

Campos, S.K., Barry, M.A., 2007. Current advances and future challenges in Adenoviral vector biology and targeting. Curr Gene Ther 7, 189-204.

Cann, K.L., Hicks, G.G., 2007. Regulation of the cellular DNA double-strand break response. Biochem Cell Biol 85, 663-674.

Capecchi, M.R., 2001. Generating mice with targeted mutations. Nat Med 7, 1086-1090.

Carbery, I.D., Ji, D., Harrington, A., Brown, V., Weinstein, E.J., Liaw, L., Cui, X., 2010. Targeted genome modification in mice using zinc-finger nucleases. Genetics 186, 451-459.

Carroll, D., 2008. Progress and prospects: zinc-finger nucleases as gene therapy agents. Gene Ther 15, 1463-1468.

Carroll, D., 2011. Genome engineering with zinc-finger nucleases. Genetics 188, 773-782.

Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Bougneres, P., Schmidt, M., Kalle, C.V., Fischer, A., Cavazzana-Calvo, M., Aubourg, P., 2012. Lentiviral hematopoietic cell gene therapy for X-linked adrenoleukodystrophy. Methods Enzymol 507, 187-198.

Case, S.S., Price, M.A., Jordan, C.T., Yu, X.J., Wang, L., Bauer, G., Haas, D.L., Xu, D., Stripecke, R., Naldini, L., Kohn, D.B., Crooks, G.M., 1999. Stable transduction of quiescent CD34(+)CD38(-) human hematopoietic cells by HIV-1-based lentiviral vectors. Proc Natl Acad Sci U S A 96, 2988-2993.

Castaigne, J.G., Guo, W., Leveille, C., Charron, D., Al-Daccak, R., 2002. A CD18-dependent protein kinase C beta-mediated alternative cell death pathway of activated monocytes. Int Immunol 14, 1003-1014.

Cathomen, T., Weitzman, M.D., 2005. Gene repair: pointing the finger at genetic disease. Gene Ther 12, 1415-1416.

Cavazzana-Calvo, M., Fischer, A., 2007. Gene therapy for severe combined immunodeficiency: are we there yet? J Clin Invest 117, 1456-1465.

Chapel, H., Geha, R., Rosen, F., 2003. Primary immunodeficiency diseases: an update. Clin Exp Immunol 132, 9-15.

Cheshier, S.H., Morrison, S.J., Liao, X., Weissman, I.L., 1999. In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. Proc Natl Acad Sci U S A 96, 3120-3125.

Chick, H.E., Nowrouzi, A., Fronza, R., McDonald, R., Kane, N., Alba, R., Delles, C., Sessa, W.C., Schmidt, M., Thrasher, A., Baker, A.H., 2012. Integrase-deficient lentiviral vectors mediate efficient gene transfer to human vascular smooth muscle cells with minimal genotoxic risk. Hum Gene Ther.

Choulika, A., Perrin, A., Dujon, B., Nicolas, J.F., 1995a. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. Mol Cell Biol 15, 1968-1973.

Choulika, A., Perrin, A., Dujon, B., Nicolas, J.F., 1995b. Induction of homologous recombination in mammalian chromosomes by using the I-*Sce*I system of *Saccharomyces cerevisiae*. Mol Cell Biol 15, 1968-1973.

Ciuffi, A., 2008. Mechanisms governing lentivirus integration site selection. Curr Gene Ther 8,6, 419-29.

Cockrell, A.S., Kafri, T., 2007. Gene delivery by lentivirus vectors. Mol Biotechnol 36, 184-204.

Coffin, J.M., Hughes, S.H., Varmus, H.E., 1997. The Interactions of Retroviruses and their Hosts.

Connelly, J.P., Barker, J.C., Pruett-Miller, S., Porteus, M.H., 2010. Gene correction by homologous recombination with zinc finger nucleases in primary cells from a mouse model of a generic recessive genetic disease. Mol Ther 18, 1103-1110.

Cornu, T.I., Cathomen, T., 2007. Targeted genome modifications using integrase-deficient lentiviral vectors. Mol Ther 15, 2107-2113.

Cornu, T.I., Thibodeau-Beganny, S., Guhl, E., Alwin, S., Eichtinger, M., Joung, J.K., Cathomen, T., 2008. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. Mol Ther 16, 352-358.

de Campos-Nebel, M., Larripa, I., Gonzalez-Cid, M., 2008. Non-homologous end joining is the responsible pathway for the repair of fludarabine-induced DNA double strand breaks in mammalian cells. Mutat Res 646, 8-16.

de Pater, S., Neuteboom, L.W., Pinas, J.E., Hooykaas, P.J., van der Zaal, B.J., 2009. ZFN-induced mutagenesis and gene-targeting in Arabidopsis through Agrobacterium-mediated floral dip transformation. Plant Biotechnol J 7, 821-835.

De Silva, I.U., McHugh, P.J., Clingen, P.H., Hartley, J.A., 2000. Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. Mol Cell Biol 20, 7980-7990.

Delenda, C., 2004. Lentiviral vectors: optimization of packaging, transduction and gene expression. J Gene Med 6 Suppl 1, S125-138.

Demaison, C., Parsley, K., Brouns, G., Scherr, M., Battmer, K., Kinnon, C., Grez, M., Thrasher, A.J., 2002. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Hum Gene Ther 13, 803-813.

Deng, C., Capecchi, M.R., 1992. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol Cell Biol 12, 3365-3371.

Dolan, P., 2006. Targeted gene repair using engineered zinc finger nuclease. MMG Basic Biotechnol eJornal 445, 7-13.

Douglas, J.T., 2004. Adenovirus-mediated gene delivery: an overview. Methods Mol Biol 246, 3-14.

Doyon, Y., Choi, V.M., Xia, D.F., Vo, T.D., Gregory, P.D., Holmes, M.C., 2010. Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. Nat Methods 7, 459-460.

Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Amacher, S.L., 2008. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol 26, 702-708.

Duan, X., Gimble, F.S., Quiocho, F.A., 1997. Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity. Cell 89, 555-564.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., Naldini, L., 1998. A third-generation lentivirus vector with a conditional packaging system. J Virol 72, 8463-8471.

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M.H., Chandrasegaran, S., 2005. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res 33, 5978-5990.

Dykstra, B., Kent, D., Bowie, M., McCaffrey, L., Hamilton, M., Lyons, K., Lee, S.J., Brinkman, R., Eaves, C., 2007. Long-term propagation of distinct hematopoietic differentiation programs in vivo. Cell Stem Cell 1, 218-229.

Ehrhardt, A., Kay, M.A., 2005. Gutted adenovirus: a rising star on the horizon? Gene Ther 12, 1540-1541.

Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J.A., Jasin, M., 1998. Gene conversion tracts from double-strand break repair in mammalian cells. Mol Cell Biol 18, 93-101.

Ellis, B.L., Hirsch, M.L., Porter, S.N., Samulski, R.J., Porteus, M.H., 2012. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. Gene Ther.

Ema, H., Morita, Y., Yamazaki, S., Matsubara, A., Seita, J., Tadokoro, Y., Kondo, H., Takano, H., Nakauchi, H., 2006. Adult mouse hematopoietic stem cells: purification and single-cell assays. Nat Protoc 1, 2979-2987.

Engelman, A., 1999. In vivo analysis of retroviral integrase structure and function. Adv Virus Res 52, 411-426.

Engelman, A., Cherepanov, P., 2008. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. PLoS Pathog 4, e1000046.

Engelman, A., Englund, G., Orenstein, J.M., Martin, M.A., Craigie, R., 1995. Multiple effects of mutations in human immunodeficiency virus type 1 integrase on viral replication. J Virol 69, 2729-2736.

Engelman, A., Mizuuchi, K., Craigie, R., 1991. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67, 1211-1221.

Epinat, J.C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., Patin, A., Zanghellini, A., Paques, F., Lacroix, E., 2003. A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. Nucleic Acids Res 31, 2952-2962.

Fankhauser, C., Izaurralde, E., Adachi, Y., Wingfield, P., Laemmli, U.K., 1991. Specific complex of human immunodeficiency virus type 1 rev and nucleolar B23 proteins: dissociation by the Rev response element. Mol Cell Biol 11, 2567-2575.

Farnet, C.M., Haseltine, W.A., 1991. Circularization of human immunodeficiency virus type 1 DNA in vitro. J Virol 65, 6942-6952.

Fernandes, R.A., Shore, D.A., Vuong, M.T., Yu, C., Zhu, X., Pereira-Lopes, S., Brouwer, H., Fennelly, J.A., Jessup, C.M., Evans, E.J., Wilson, I.A., Davis, S.J., 2012. T cell receptors are structures capable of initiating signaling in the absence of large conformational rearrangements. J Biol Chem 287, 13324-13335.

Fernex, C., Dubreuil, P., Mannoni, P., Bagnis, C., 1997. Cre/loxP-mediated excision of a neomycin resistance expression unit from an integrated retroviral vector increases long terminal repeat-driven transcription in human hematopoietic cells. J Virol 71, 7533-7540.

Fischer, A., Cavazzana-Calvo, M., 2005. Integration of retroviruses: a fine balance between efficiency and danger. PLoS Med 2, e10.

Fischer, A., Hacein-Bey-Abina, S., Cavazzana-Calvo, M., 2012. Strategies for retrovirus-based correction of severe, combined immunodeficiency (SCID). Methods Enzymol 507, 15-27.

Foley, J.E., Yeh, J.R., Maeder, M.L., Reyon, D., Sander, J.D., Peterson, R.T., Joung, J.K., 2009. Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). PLoS One 4, e4348.

Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M., Naldini, L., 2000. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. Nat Genet 25, 217-222.

Fulop, G.M., Phillips, R.A., 1986. Full reconstitution of the immune deficiency in scid mice with normal stem cells requires low-dose irradiation of the recipients. J Immunol 136, 4438-4443.

Gabriel, R., Lombardo, A., Arens, A., Miller, J.C., Genovese, P., Kaeppel, C., Nowrouzi, A., Bartholomae, C.C., Wang, J., Friedman, G., Holmes, M.C., Gregory, P.D., Glimm, H., Schmidt, M., Naldini, L., von Kalle, C., 2011. An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol 29, 816-823.

Gaj, T., Guo, J., Kato, Y., Sirk, S.J., Barbas, C.F., 3rd, 2012. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods 9, 805-807.

Gaspar, H.B., Cooray, S., Gilmour, K.C., Parsley, K.L., Adams, S., Howe, S.J., Al Ghonaium, A., Bayford, J., Brown, L., Davies, E.G., Kinnon, C., Thrasher, A.J., 2011a. Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. Sci Transl Med 3, 97ra79.

Gaspar, H.B., Cooray, S., Gilmour, K.C., Parsley, K.L., Zhang, F., Adams, S., Bjorkegren, E., Bayford, J., Brown, L., Davies, E.G., Veys, P., Fairbanks, L., Bordon, V., Petropolou, T., Kinnon, C., Thrasher, A.J., 2011b. Hematopoietic stem cell gene therapy for adenosine deaminase-deficient severe combined immunodeficiency leads to long-term immunological recovery and metabolic correction. Sci Transl Med 3, 97ra80.

Gellhaus, K., Cornu, T.I., Heilbronn, R., Cathomen, T., 2010. Fate of recombinant adeno-associated viral vector genomes during DNA double-strand break-induced gene targeting in human cells. Hum Gene Ther 21, 543-553.

Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood, A., Cui, X., Meng, X., Vincent, A., Lam, S., Michalkiewicz, M., Schilling, R., Foeckler, J., Kalloway, S., Weiler, H., Menoret, S., Anegon, I., Davis, G.D., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Jacob, H.J., Buelow, R., 2009a. Knockout rats via embryo microinjection of zinc-finger nucleases. Science 325, 433.

Geurts, A.M., Cost, G.J., Remy, S., Cui, X., Tesson, L., Usal, C., Menoret, S., Jacob, H.J., Anegon, I., Buelow, R., 2009b. Generation of gene-specific mutated rats using zinc-finger nucleases. Methods Mol Biol 597, 211-225.

Gilley, D., Tanaka, H., Hande, M.P., Kurimasa, A., Li, G.C., Oshimura, M., Chen, D.J., 2001. DNA-PKcs is critical for telomere capping. Proc Natl Acad Sci U S A 98, 15084-15088.

Ginn, S.L., Liao, S.H., Dane, A.P., Hu, M., Hyman, J., Finnie, J.W., Zheng, M., Cavazzana-Calvo, M., Alexander, S.I., Thrasher, A.J., Alexander, I.E., 2010. Lymphomagenesis in SCID-X1 mice following lentivirus-mediated phenotype correction independent of insertional mutagenesis and gammac overexpression. Mol Ther 18, 965-976.

Glimm, H., Oh, I.H., Eaves, C.J., 2000. Human hematopoietic stem cells stimulated to proliferate in vitro lose engraftment potential during their S/G(2)/M transit and do not reenter G(0). Blood 96, 4185-4193.

Goff, S.P., 2001. Intracellular trafficking of retroviral genomes during the early phase of infection: viral exploitation of cellular pathways. J Gene Med 3, 517-528.

Goueli, S.a.K.H., 1997. SignaTECTTM DNA-Dependent Protein Kinase

Assay System. Promega Notes Magazine Promega Corporation, 7- 9

Grabher, C., Wittbrodt, J., 2007. Meganuclease and transposon mediated transgenesis in medaka. Genome Biol 8 Suppl 1, S10.

Graham, F.L., Smiley, J., Russell, W.C., Nairn, R., 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol 36, 59-74.

Guirouilh-Barbat, J., Huck, S., Lopez, B.S., 2008. S-phase progression stimulates both the mutagenic KU-independent pathway and mutagenic processing of KU-dependent intermediates, for nonhomologous end joining. Oncogene 27, 1726-1736.

Hande, M.P., Samper, E., Lansdorp, P., Blasco, M.A., 1999. Telomere length dynamics and chromosomal instability in cells derived from telomerase null mice. J Cell Biol 144, 589-601.

Harley, C.B., Futcher, A.B., Greider, C.W., 1990. Telomeres shorten during ageing of human fibroblasts. Nature 345, 458-460.

Harper, J.W., Elledge, S.J., 2007. The DNA damage response: ten years after. Mol Cell 28, 739-745.

Harrington, L., McPhail, T., Mar, V., Zhou, W., Oulton, R., Bass, M.B., Arruda, I., Robinson, M.O., 1997. A mammalian telomerase-associated protein. Science 275, 973-977.

Hatada, S., Nikkuni, K., Bentley, S.A., Kirby, S., Smithies, O., 2000. Gene correction in hematopoietic progenitor cells by homologous recombination. Proc Natl Acad Sci U S A 97, 13807-13811.

Heath, P.J., Stephens, K.M., Monnat, R.J., Jr., Stoddard, B.L., 1997. The structure of I-Crel, a group I intron-encoded homing endonuclease. Nat Struct Biol 4, 468-476.

Hendrie, P.C., Russell, D.W., 2005. Gene targeting with viral vectors. Mol Ther 12, 9-17.

Hida, K., Lai, S.K., Suk, J.S., Won, S.Y., Boyle, M.P., Hanes, J., 2011. Common gene therapy viral vectors do not efficiently penetrate sputum from cystic fibrosis patients. PLoS One 6, e19919.

Hinnen, A., Hicks, J.B., Fink, G.R., 1978. Transformation of yeast. Proc Natl Acad Sci U S A 75, 1929-1933.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., Meng, X., Miller, J.C., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Jaenisch, R., 2009. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol 27, 851-857.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., Zeitler, B., Cherone, J.M., Meng, X., Hinkley, S.J., Rebar, E.J., Gregory, P.D., Urnov, F.D., Jaenisch, R., 2011. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29, 731-734.

Holmquist, G.P., 1998. Endogenous lesions, S-phase-independent spontaneous mutations, and evolutionary strategies for base excision repair. Mutat Res 400, 59-68.

Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C., Cannon, P.M., 2010. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol 28, 839-847.

Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D., Gilmour, K.C., Adams, S., Thornhill, S.I., Parsley, K.L., Staal, F.J., Gale, R.E., Linch, D.C., Bayford, J., Brown, L., Quaye, M., Kinnon, C., Ancliff, P., Webb, D.K., Schmidt, M., von Kalle, C., Gaspar, H.B., Thrasher, A.J., 2008. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest 118, 3143-3150.

http://www.cellectis.com/genome-engineering/meganucleases/engineered-meganucleases/meganuclease-technologies/, accessed on 20/06/2012.

http://www.nobelprize.org/nobel_prizes/medicine/laureates/2007/press.html, accessed on 4/7/2012.

http://www.wiley.co.uk/genetherapy/clinical, accessed on 16/08/2012.

Huston, M.W., van Til, N.P., Visser, T.P., Arshad, S., Brugman, M.H., Cattoglio, C., Nowrouzi, A., Li, Y., Schambach, A., Schmidt, M., Baum, C., von Kalle, C., Mavilio, F., Zhang, F., Blundell, M.P., Thrasher, A.J., Verstegen, M.M., Wagemaker, G., 2011. Correction of murine SCID-X1 by lentiviral gene therapy using a codon-optimized IL2RG gene and minimal pretransplant conditioning. Mol Ther 19, 1867-1877.

Igoucheva, O., Alexeev, V., Yoon, K., 2004. Oligonucleotide-directed mutagenesis and targeted gene correction: a mechanistic point of view. Curr Mol Med 4, 445-463.

Iiizumi, S., Kurosawa, A., So, S., Ishii, Y., Chikaraishi, Y., Ishii, A., Koyama, H., Adachi, N., 2008. Impact of non-homologous end-joining deficiency on random and targeted DNA integration: implications for gene targeting. Nucleic Acids Res 36, 6333-6342.

Ildefonso, C.J., Bond, W.S., Al-Tawashi, A.R., Hurwitz, M.Y., Hurwitz, R.L., 2012. The liberation of CD44 intracellular domain modulates adenoviral vector transgene expression. J Biol Chem.

Johnson, R.D., Jasin, M., 2001. Double-strand-break-induced homologous recombination in mammalian cells. Biochem Soc Trans 29, 196-201.

Kandavelou, K., Mani, M., Durai, S., Chandrasegaran, S., 2005. "Magic" scissors for genome surgery. Nat Biotechnol 23, 686-687.

Kandavelou, K., Ramalingam, S., London, V., Mani, M., Wu, J., Alexeev, V., Civin, C.I., Chandrasegaran, S., 2009. Targeted manipulation of mammalian genomes using designed zinc finger nucleases. Biochem Biophys Res Commun 388, 56-61.

Kawai, J., Hirose, K., Fushiki, S., Hirotsune, S., Ozawa, N., Hara, A., Hayashizaki, Y., Watanabe, S., 1994. Comparison of DNA methylation patterns among mouse cell lines by restriction landmark genomic scanning. Mol Cell Biol 14, 7421-7427.

Kawase, Y., Ladage, D., Hajjar, R.J., 2011. Rescuing the failing heart by targeted gene transfer. J Am Coll Cardiol 57, 1169-1180.

Kay, M.A., Glorioso, J.C., Naldini, L., 2001. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. Nat Med 7, 33-40.

Khare, R., Chen, C.Y., Weaver, E.A., Barry, M.A., 2011. Advances and future challenges in adenoviral vector pharmacology and targeting. Curr Gene Ther 11, 241-258.

Kirchgessner, C.U., Patil, C.K., Evans, J.W., Cuomo, C.A., Fried, L.M., Carter, T., Oettinger, M.A., Brown, J.M., 1995. DNA-dependent kinase (p350) as a candidate gene for the murine SCID defect. Science 267, 1178-1183.

Klug, A., 2005. The discovery of zinc fingers nucleases and their development for practical applications in gene regulation. Proc Japan Acad 81 87-102.

Kohn, D.B., 2001. Gene therapy for genetic haematological disorders and immunodeficiencies. J Intern Med 249, 379-390.

Kohn, D.B., 2010. Update on gene therapy for immunodeficiencies. Clin Immunol 135, 247-254.

Kotnis, A., Du, L., Liu, C., Popov, S.W., Pan-Hammarstrom, Q., 2009. Non-homologous end joining in class switch recombination: the beginning of the end. Philos Trans R Soc Lond B Biol Sci 364, 653-665.

Kung, S.K., An, D.S., Chen, I.S., 2000. A murine leukemia virus (MuLV) long terminal repeat derived from rhesus macaques in the context of a lentivirus vector and MuLV gag sequence results in high-level gene expression in human T lymphocytes. J Virol 74, 3668-3681.

Kutner, R.H., Zhang, X.Y., Reiser, J., 2009. Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. Nat Protoc 4, 495-505.

Lai, L., Alaverdi, N., Maltais, L., Morse, H.C., 3rd, 1998. Mouse cell surface antigens: nomenclature and immunophenotyping. J Immunol 160, 3861-3868.

Langford-Smith, A., Wilkinson, F.L., Langford-Smith, K.J., Holley, R.J., Sergijenko, A., Howe, S.J., Bennett, W.R., Jones, S.A., Wraith, J., Merry, C.L., Wynn, R.F., Bigger, B.W., 2012. Hematopoietic Stem Cell and Gene Therapy Corrects Primary Neuropathology and Behavior in Mucopolysaccharidosis IIIA Mice. Mol Ther 20, 1610-1621.

Leavitt, A.D., Robles, G., Alesandro, N., Varmus, H.E., 1996. Human immunodeficiency virus type 1 integrase mutants retain in vitro integrase activity yet fail to integrate viral DNA efficiently during infection. J Virol 70, 721-728.

Leavitt, A.D., Shiue, L., Varmus, H.E., 1993. Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. J Biol Chem 268, 2113-2119.

Leblond, V., Autran, B., Cesbron, J.Y., 1997. The SCID mouse mutant: definition and potential use as a model for immune and hematological disorders. Hematol Cell Ther 39, 213-221.

Lehmusvaara S. , Meriläinen O., Hakkarainen T. , Wahlfors J., 2005. Enhancement of Adeno- and Lentivirus Mediated Gene Transfer in Human Tumor Cell Lines by Cell Permeable Peptides and Polycations Molecular Therapy 11, S189.

Lesch, H.P., Laitinen, A., Peixoto, C., Vicente, T., Makkonen, K.E., Laitinen, L., Pikkarainen, J.T., Samaranayake, H., Alves, P.M., Carrondo, M.J., Yla-Herttuala, S., Airenne, K.J., 2011. Production and purification of lentiviral vectors generated in 293T suspension cells with baculoviral vectors. Gene Ther 18, 531-538.

Levine, B.L., Humeau, L.M., Boyer, J., MacGregor, R.R., Rebello, T., Lu, X., Binder, G.K., Slepushkin, V., Lemiale, F., Mascola, J.R., Bushman, F.D., Dropulic, B., June, C.H., 2006. Gene transfer in humans using a conditionally replicating lentiviral vector. Proc Natl Acad Sci U S A 103, 17372-17377.

Li, Haurigot, V., Doyon, Y., Li, T., Wong, S.Y., Bhagwat, A.S., Malani, N., Anguela, X.M., Sharma, R., Ivanciu, L., Murphy, S.L., Finn, J.D., Khazi, F.R., Zhou, S., Paschon, D.E., Rebar, E.J., Bushman, F.D., Gregory, P.D., Holmes, M.C., High, K.A., 2011a. In vivo genome editing restores haemostasis in a mouse model of haemophilia. Nature 475, 217-221.

Li, Husic, N., Lin, Y., Christensen, H., Malik, I., McIver, S., LaPash Daniels, C.M., Harris, D.A., Kotzbauer, P.T., Goldberg, M.P., Snider, B.J., 2010. Optimal promoter

usage for lentiviral vector-mediated transduction of cultured central nervous system cells. J Neurosci Methods 189, 56-64.

Li, Liu, B., Spalding, M.H., Weeks, D.P., Yang, B., 2012a. High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol 30, 390-392.

Li, Sun, L., Xiao, L., Liu, F.Y., 2012b. Gene delivery in peritoneal dialysis related peritoneal fibrosis research. Chin Med J (Engl) 125, 2219-2224.

Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S.Y., Bhagwat, A.S., Malani, N., Anguela, X.M., Sharma, R., Ivanciu, L., Murphy, S.L., Finn, J.D., Khazi, F.R., Zhou, S., Paschon, D.E., Rebar, E.J., Bushman, F.D., Gregory, P.D., Holmes, M.C., High, K.A., 2011b. In vivo genome editing restores haemostasis in a mouse model of haemophilia. Nature 475, 217-221.

Lieber, M.R., Ma, Y., Pannicke, U., Schwarz, K., 2003. Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol 4, 712-720.

Lim, K.I., 2012. Retroviral integration profiles: their determinants and implications for gene therapy. BMB Rep 45, 207-212.

Linemeyer, D.L., Menke, J.G., Ruscetti, S.K., Evans, L.H., Scolnick, E.M., 1982. Envelope gene sequences which encode the gp52 protein of spleen focus-forming virus are required for the induction of erythroid cell proliferation. J Virol 43, 223-233.

Liu, P.Q., Chan, E.M., Cost, G.J., Zhang, L., Wang, J., Miller, J.C., Guschin, D.Y., Reik, A., Holmes, M.C., Mott, J.E., Collingwood, T.N., Gregory, P.D., 2010. Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. Biotechnol Bioeng 106, 97-105.

Logan, A.C., Haas, D.L., Kafri, T., Kohn, D.B., 2004. Integrated self-inactivating lentiviral vectors produce full-length genomic transcripts competent for encapsidation and integration. J Virol 78, 8421-8436.

Logan, A.C., Lutzko, C., Kohn, D.B., 2002. Advances in lentiviral vector design for gene-modification of hematopoietic stem cells. Curr Opin Biotechnol 13, 429-436.

Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., Holmes, M.C., Naldini, L., 2007. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol 25, 1298-1306.

Loong, S.L., Korzh, S., Price, A., 2004. Reduced DNA-dependent protein kinase activity in two cell lines derived from adult cancer patients with late radionecrosis. Oncogene 23, 5562-5566.

Lu, R., Limon, A., Devroe, E., Silver, P.A., Cherepanov, P., Engelman, A., 2004. Class II integrase mutants with changes in putative nuclear localization signals are primarily blocked at a postnuclear entry step of human immunodeficiency virus type 1 replication. J Virol 78, 12735-12746.

Macville, M., Schrock, E., Padilla-Nash, H., Keck, C., Ghadimi, B.M., Zimonjic, D., Popescu, N., Ried, T., 1999. Comprehensive and definitive molecular cytogenetic characterization of HeLa cells by spectral karyotyping. Cancer Res 59, 141-150.

Mahfouz, M.M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., Zhu, J.K., 2011. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci U S A 108, 2623-2628.

Maldonado-Lopez, R., De Smedt, T., Pajak, B., Heirman, C., Thielemans, K., Leo, O., Urbain, J., Maliszewski, C.R., Moser, M., 1999. Role of CD8alpha+ and CD8alpha- dendritic cells in the induction of primary immune responses in vivo. J Leukoc Biol 66, 242-246.

Malphettes, L., Freyvert, Y., Chang, J., Liu, P.Q., Chan, E., Miller, J.C., Zhou, Z., Nguyen, T., Tsai, C., Snowden, A.W., Collingwood, T.N., Gregory, P.D., Cost, G.J., 2010. Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. Biotechnol Bioeng 106, 774-783.

Mani, M., Smith, J., Kandavelou, K., Berg, J.M., Chandrasegaran, S., 2005. Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage. Biochem Biophys Res Commun 334, 1191-1197.

Marras, S.A., 2006. Selection of fluorophore and quencher pairs for fluorescent nucleic acid hybridization probes. Methods Mol Biol 335, 3-16.

Marton, I., Zuker, A., Shklarman, E., Zeevi, V., Tovkach, A., Roffe, S., Ovadis, M., Tzfira, T., Vainstein, A., 2010. Nontransgenic genome modification in plant cells. Plant Physiol 154, 1079-1087.

Mashimo, T., Takizawa, A., Voigt, B., Yoshimi, K., Hiai, H., Kuramoto, T., Serikawa, T., 2010. Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. PLoS One 5, e8870.

Matsuda, M., Arai, A., Nakamura, Y., Fujisawa, R., Masuda, M., 2009. Host cell-specific effects of lentiviral accessory proteins on the eukaryotic cell cycle progression. Microbes Infect 11, 646-653.

Matsuzaki, Y., Kinjo, K., Mulligan, R.C., Okano, H., 2004. Unexpectedly efficient homing capacity of purified murine hematopoietic stem cells. Immunity 20, 87-93.

Mauro, M., Rego, M.A., Boisvert, R.A., Esashi, F., Cavallo, F., Jasin, M., Howlett, N.G., 2012. p21 promotes error-free replication-coupled DNA double-strand break repair. Nucleic Acids Res 40, 8348-8360.

McKinnon, P.J., Caldecott, K.W., 2007. DNA strand break repair and human genetic disease. Annu Rev Genomics Hum Genet 8, 37-55.

Medvinsky, A., Rybtsov, S., Taoudi, S., 2011. Embryonic origin of the adult hematopoietic system: advances and questions. Development 138, 1017-1031.

Micklem, H.S., Ford, C.E., Evans, E.P., Ogden, D.A., Papworth, D.S., 1972. Competitive in vivo proliferation of foetal and adult haematopoietic cells in lethally irradiated mice. J Cell Physiol 79, 293-298.

Mikkola, H., Woods, N.B., Sjogren, M., Helgadottir, H., Hamaguchi, I., Jacobsen, S.E., Trono, D., Karlsson, S., 2000. Lentivirus gene transfer in murine hematopoietic progenitor cells is compromised by a delay in proviral integration and results in transduction mosaicism and heterogeneous gene expression in progeny cells. J Virol 74, 11911-11918.

Miller, J.C., Holmes, M.C., Wang, J., Guschin, D.Y., Lee, Y.L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Kim, K.A., Gregory, P.D., Pabo, C.O., Rebar, E.J., 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25, 778-785.

Miyoshi, H., Blomer, U., Takahashi, M., Gage, F.H., Verma, I.M., 1998. Development of a self-inactivating lentivirus vector. J Virol 72, 8150-8157.

Modell, V., Gee, B., Lewis, D.B., Orange, J.S., Roifman, C.M., Routes, J.M., Sorensen, R.U., Notarangelo, L.D., Modell, F., 2011. Global study of primary immunodeficiency diseases (PI)--diagnosis, treatment, and economic impact: an updated report from the Jeffrey Modell Foundation. Immunol Res 51, 61-70.

Modlich, U., Navarro, S., Zychlinski, D., Maetzig, T., Knoess, S., Brugman, M.H., Schambach, A., Charrier, S., Galy, A., Thrasher, A.J., Bueren, J., Baum, C., 2009. Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. Mol Ther 17, 1919-1928.

Moehle, E.A., Rock, J.M., Lee, Y.L., Jouvenot, Y., DeKelver, R.C., Gregory, P.D., Urnov, F.D., Holmes, M.C., 2007. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci U S A 104, 3055-3060.

Mohty, M., Apperley, J.F., 2010. Long-term physiological side effects after allogeneic bone marrow transplantation. Hematology Am Soc Hematol Educ Program 2010, 229-236.

Moldt, B., Staunstrup, N.H., Jakobsen, M., Yanez-Munoz, R.J., Mikkelsen, J.G., 2008. Genomic insertion of lentiviral DNA circles directed by the yeast Flp recombinase. BMC Biotechnol 8, 60.

Mondorf, S., Deppenmeier, U., Welte, C., 2012. A Novel Inducible Protein Production System and Neomycin Resistance as Selection Marker for Methanosarcina mazei. Archaea 2012, 973743.

Monteith, C.E., Chelack, B.J., Davis, W.C., Haines, D.M., 1996. Identification of monoclonal antibodies for immunohistochemical staining of feline B lymphocytes in frozen and formalin-fixed paraffin-embedded tissues. Can J Vet Res 60, 193-198.

Morozova, O., Hirst, M., Marra, M.A., 2009. Applications of new sequencing technologies for transcriptome analysis. Annu Rev Genomics Hum Genet 10, 135-151.

Montini, E.,Cesana, D.,Schmidt, M.,Sanvito, F.,Ponzoni, M.,Bartholomae, C.,Sergi Sergi, L.,Benedicenti, F.,Ambrosi, A.,Di Serio, C.,Doglioni, C.,von Kalle, C.,Naldini, L., 2006. Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat Biotechnol 24,6, 687-96.

Moscou, M.J., Bogdanove, A.J., 2009. A simple cipher governs DNA recognition by TAL effectors. Science 326, 1501.

Moshous, D., Callebaut, I., de Chasseval, R., Corneo, B., Cavazzana-Calvo, M., Le Deist, F., Tezcan, I., Sanal, O., Bertrand, Y., Philippe, N., Fischer, A., de Villartay, J.P., 2001. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. Cell 105, 177-186.

Mostoslavsky, G., Fabian, A.J., Rooney, S., Alt, F.W., Mulligan, R.C., 2006. Complete correction of murine Artemis immunodeficiency by lentiviral vector-mediated gene transfer. Proc Natl Acad Sci U S A 103, 16406-16411.

Moure, C.M., Gimble, F.S., Quiocho, F.A., 2002. Crystal structure of the intein homing endonuclease PI-SceI bound to its recognition sequence. Nat Struct Biol 9, 764-770.

Moure, C.M., Gimble, F.S., Quiocho, F.A., 2003. The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. J Mol Biol 334, 685-695.

Muller, C., Calsou, P., Salles, B., 2000. The activity of the DNA-dependent protein kinase (DNA-PK) complex is determinant in the cellular response to nitrogen mustards. Biochimie 82, 25-28.

Mulligan, R.C., 1993. The basic science of gene therapy. Science 260, 926-932.

Mussolino, C., Morbitzer, R., Lutge, F., Dannemann, N., Lahaye, T., Cathomen, T., 2011. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res 39, 9283-9293.

Mussolino, C. and Cathomen, T., 2012. TALE nucleases: tailored genome engineering made easy. Curr Opin Biotechnol, 23, 5, 644-50.

Nakajima, N., Lu, R., Engelman, A., 2001. Human immunodeficiency virus type 1 replication in the absence of integrase-mediated dna recombination: definition of permissive and nonpermissive T-cell lines. J Virol 75, 7944-7955.

Nakayama, M., 2010. Homologous recombination in human iPS and ES cells for use in gene correction therapy. Drug Discov Today 15, 198-202.

Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., Trono, D., 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science 272, 263-267.

Negri, D.R., Michelini, Z., Baroncelli, S., Spada, M., Vendetti, S., Bona, R., Leone, P., Klotman, M.E., Cara, A., 2010. Nonintegrating Lentiviral Vector-Based Vaccine Efficiently Induces Functional and Persistent CD8+ T Cell Responses in Mice. J Biomed Biotechnol 2010, 534501.

Nemerow, G.R., Pache, L., Reddy, V., Stewart, P.L., 2009. Insights into adenovirus host cell interactions from structural studies. Virology 384, 380-388.

Niedernhofer, L.J., 2008. DNA repair is crucial for maintaining hematopoietic stem cell function. DNA Repair (Amst) 7, 523-529.

Niehues, T., Perez-Becker, R., Schuetz, C., 2010. More than just SCID--the phenotypic range of combined immunodeficiencies associated with mutations in the recombinase activating genes (RAG) 1 and 2. Clin Immunol 135, 183-192.

Nielsen, T.T., Jakobsson, J., Rosenqvist, N., Lundberg, C., 2009. Incorporating double copies of a chromatin insulator into lentiviral vectors results in less viral integrants. BMC Biotechnol 9, 13.

Nightingale, S.J., Hollis, R.P., Pepper, K.A., Petersen, D., Yu, X.J., Yang, C., Bahner, I., Kohn, D.B., 2006. Transient gene expression by nonintegrating lentiviral vectors. Mol Ther 13, 1121-1132.

Notarangelo, L.D., 2010. Primary immunodeficiencies. J Allergy Clin Immunol 125, S182-194.

Notarangelo, L.D., Fischer, A., Geha, R.S., Casanova, J.L., Chapel, H., Conley, M.E., Cunningham-Rundles, C., Etzioni, A., Hammartrom, L., Nonoyama, S., Ochs, H.D., Puck, J., Roifman, C., Seger, R., Wedgwood, J., 2009. Primary immunodeficiencies: 2009 update. J Allergy Clin Immunol 124, 1161-1178.

Okada, Y., Ueshin, Y., Hasuwa, H., Takumi, K., Okabe, M., Ikawa, M., 2009. Targeted gene modification in mouse ES cells using integrase-defective lentiviral vectors. Genesis.

Oleykowski, C.A., Bronson Mullins, C.R., Godwin, A.K., Yeung, A.T., 1998. Mutation detection using a novel plant endonuclease. Nucleic Acids Res 26, 4597-4602.

Olsen, P.A., Gelazauskaite, M., Randol, M., Krauss, S., 2010. Analysis of illegitimate genomic integration mediated by zinc-finger nucleases: implications for specificity of targeted gene correction. BMC Mol Biol 11, 35.

Olsen, P.A., Randol, M., Luna, L., Brown, T., Krauss, S., 2005. Genomic sequence correction by single-stranded DNA oligonucleotides: role of DNA synthesis and chemical modifications of the oligonucleotide ends. J Gene Med 7, 1534-1544.

Olsen, P.A., Solhaug, A., Booth, J.A., Gelazauskaite, M., Krauss, S., 2009. Cellular responses to targeted genomic sequence modification using single-stranded oligonucleotides and zinc-finger nucleases. DNA Repair (Amst) 8, 298-308.

Orlando, S.J., Santiago, Y., DeKelver, R.C., Freyvert, Y., Boydston, E.A., Moehle, E.A., Choi, V.M., Gopalan, S.M., Lou, J.F., Li, J., Miller, J.C., Holmes, M.C., Gregory, P.D., Urnov, F.D., Cost, G.J., 2010. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res 38, e152.

Orlic, D., 2002. Stem cell repair in ischemic heart disease: an experimental model. Int J Hematol 76 Suppl 1, 144-145.

Orlowski, J., Boniecki, M., Bujnicki, J.M., 2007. I-Ssp6803I: the first homing endonuclease from the PD-(D/E)XK superfamily exhibits an unusual mode of DNA recognition. Bioinformatics 23, 527-530.

Osakabe, K., Osakabe, Y., Toki, S., 2010. Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases. Proc Natl Acad Sci U S A.

Osawa, M., Hanada, K., Hamada, H., Nakauchi, H., 1996. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. Science 273, 242-245.

Papworth, M., Moore, M., Isalan, M., Minczuk, M., Choo, Y., Klug, A., 2003. Inhibition of herpes simplex virus 1 gene expression by designer zinc-finger transcription factors. Proc Natl Acad Sci U S A 100, 1621-1626.

Pattanayak, V., Ramirez, C.L., Joung, J.K., Liu, D.R., 2011. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods 8, 765-770.

Paulovicova, E., Korcova, J., Farkas, P., Bystricky, S., 2010. Immunological efficacy of glycoconjugates derived from Vibrio cholerae O1 serotype Ogawa detoxified LPS in mice. J Med Microbiol 59, 1440-1448.

Pauza, C.D., 1990. Two bases are deleted from the termini of HIV-1 linear DNA during integrative recombination. Virology 179, 886-889.

Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.L., Guschin, D.Y., Rupniewski, I., Waite, A.J., Carpenito, C., Carroll, R.G., Orange, J.S., Urnov, F.D., Rebar, E.J., Ando, D., Gregory, P.D., Riley, J.L., Holmes, M.C., June, C.H., 2008. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 26, 808-816.

Philippe, S., Sarkis, C., Barkats, M., Mammeri, H., Ladroue, C., Petit, C., Mallet, J., Serguera, C., 2006. Lentiviral vectors with a defective integrase allow efficient and sustained transgene expression in vitro and in vivo. Proc Natl Acad Sci U S A 103, 17684-17689.

Philpott, N.J., Thrasher, A.J., 2007. Use of nonintegrating lentiviral vectors for gene therapy. Hum Gene Ther 18, 483-489.

Piechaczek, C., Fetzer, C., Baiker, A., Bode, J., Lipps, H.J., 1999. A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. Nucleic Acids Res 27, 426-428.

Porteus, M.H., 2009. Plant biotechnology: Zinc fingers on target. Nature 459, 337-338.

Porteus, M.H., Baltimore, D., 2003. Chimeric nucleases stimulate gene targeting in human cells. Science 300, 763.

Porteus, M.H., Carroll, D., 2005. Gene targeting using zinc finger nucleases. Nat Biotechnol 23, 967-973.

Priestley, A., Beamish, H.J., Gell, D., Amatucci, A.G., Muhlmann-Diaz, M.C., Singleton, B.K., Smith, G.C., Blunt, T., Schalkwyk, L.C., Bedford, J.S., Jackson, S.P., Jeggo, P.A., Taccioli, G.E., 1998. Molecular and biochemical characterisation of DNA-dependent protein kinase-defective rodent mutant irs-20. Nucleic Acids Res 26, 1965-1973.

Pruett-Miller, S.M., Connelly, J.P., Maeder, M.L., Joung, J.K., Porteus, M.H., 2008. Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. Mol Ther 16, 707-717.

Pruett-Miller, S.M., Reading, D.W., Porter, S.N., Porteus, M.H., 2009. Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. PLoS Genet 5, e1000376.

Qasim, W., Gaspar, H.B., Thrasher, A.J., 2009. Progress and prospects: gene therapy for inherited immunodeficiencies. Gene Ther 16, 1285-1291.

Qian, H., Buza-Vidas, N., Hyland, C.D., Jensen, C.T., Antonchuk, J., Mansson, R., Thoren, L.A., Ekblom, M., Alexander, W.S., Jacobsen, S.E., 2007. Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. Cell Stem Cell 1, 671-684.

Qiu, P., Shandilya, H., D'Alessio, J.M., O'Connor, K., Durocher, J., Gerard, G.F., 2004. Mutation detection using Surveyor nuclease. Biotechniques 36, 702-707.

Radecke, S., Radecke, F., Peter, I., Schwarz, K., 2006. Physical incorporation of a single-stranded oligodeoxynucleotide during targeted repair of a human chromosomal locus. J Gene Med 8, 217-228.

Ramirez, C.L., Certo, M.T., Mussolino, C., Goodwin, M.J., Cradick, T.J., McCaffrey, A.P., Cathomen, T., Scharenberg, A.M., Joung, J.K., 2012. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. Nucleic Acids Res 40, 5560-5568.

Raus, S., Coin, S., Monsurro, V., 2011. Adenovirus as a new agent for multiple myeloma therapies: Opportunities and restrictions. Korean J Hematol 46, 229-238.

Remy, S., Tesson, L., Menoret, S., Usal, C., Scharenberg, A.M., Anegon, I., 2009. Zinc-finger nucleases: a powerful tool for genetic engineering of animals. Transgenic Res 19, 363-371.

Reynolds, L., Ullman, C., Moore, M., Isalan, M., West, M.J., Clapham, P., Klug, A., Choo, Y., 2003. Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. Proc Natl Acad Sci U S A 100, 1615-1620.

Rivat, C., Santilli, G., Gaspar, H.B., Thrasher, A.J., 2012. Gene therapy for primary immunodeficiencies. Hum Gene Ther 23, 668-675.

Rivera-Munoz, P., Soulas-Sprauel, P., Le Guyader, G., Abramowski, V., Bruneau, S., Fischer, A., Paques, F., de Villartay, J.P., 2009. Reduced immunoglobulin class switch recombination in the absence of Artemis. Blood 114, 3601-3609.

Rooney, S., Sekiguchi, J., Zhu, C., Cheng, H.L., Manis, J., Whitlow, S., DeVido, J., Foy, D., Chaudhuri, J., Lombard, D., Alt, F.W., 2002. Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. Mol Cell 10, 1379-1390.

Rosendaal, M., Hodgson, G.S., Bradley, T.R., 1979. Organization of haemopoietic stem cells: the generation-age hypothesis. Cell Tissue Kinet 12, 17-29.

Rusche, L.N., Kirchmaier, A.L., Rine, J., 2003. The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae. Annu Rev Biochem 72, 481-516.

Samlowski, W.E., Daynes, R.A., 1985. Bone marrow engraftment efficiency is enhanced by competitive inhibition of the hepatic asialoglycoprotein receptor. Proc Natl Acad Sci U S A 82, 2508-2512.

San Filippo, J., Sung, P., Klein, H., 2008. Mechanism of eukaryotic homologous recombination. Annu Rev Biochem 77, 229-257.

Sanchez-Puig, J.M., Blasco, R., 2000. Puromycin resistance (pac) gene as a selectable marker in vaccinia virus. Gene 257, 57-65.

Sander, J.D., Cade, L., Khayter, C., Reyon, D., Peterson, R.T., Joung, J.K., Yeh, J.R., 2011. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29, 697-698.

Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., Dobbs, D., 2010. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res 38, W462-468.

Sandovici, M., Deelman, L.E., Smit-van Oosten, A., van Goor, H., Rots, M.G., de Zeeuw, D., Henning, R.H., 2006. Enhanced transduction of fibroblasts in transplanted kidney with an adenovirus having an RGD motif in the HI loop. Kidney Int 69, 45-52.

Santiago, Y., Chan, E., Liu, P.Q., Orlando, S., Zhang, L., Urnov, F.D., Holmes, M.C., Guschin, D., Waite, A., Miller, J.C., Rebar, E.J., Gregory, P.D., Klug, A., Collingwood, T.N., 2008. Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A 105, 5809-5814.

Schambach, A., Bohne, J., Baum, C., Hermann, F.G., Egerer, L., von Laer, D., Giroglou, T., 2006. Woodchuck hepatitis virus post-transcriptional regulatory element deleted from X protein and promoter sequences enhances retroviral vector titer and expression. Gene Ther 13, 641-645.

Scherer, W.F., 1954. Studies on the propagation in vitro of poliomyelitis viruses. VI. Effect on virus yield of cell population, virus inoculum and temperature of incubation. J Immunol 73, 331-336.

Schreiner, S., Martinez, R., Groitl, P., Rayne, F., Vaillant, R., Wimmer, P., Bossis, G., Sternsdorf, T., Marcinowski, L., Ruzsics, Z., Dobner, T., Wodrich, H., 2012. Transcriptional activation of the adenoviral genome is mediated by capsid protein VI. PLoS Pathog 8, e1002549.

Schroeder, T., 2010. Hematopoietic stem cell heterogeneity: subtypes, not unpredictable behavior. Cell Stem Cell 6, 203-207.

Shepherd, B.E., Kiem, H.P., Lansdorp, P.M., Dunbar, C.E., Aubert, G., LaRochelle, A., Seggewiss, R., Guttorp, P., Abkowitz, J.L., 2007. Hematopoietic stem-cell behavior in nonhuman primates. Blood 110, 1806-1813.

Shida, H., 2012. Role of Nucleocytoplasmic RNA Transport during the Life Cycle of Retroviruses. Front Microbiol 3, 179.

Shimizu, Y., Sollu, C., Meckler, J.F., Adriaenssens, A., Zykovich, A., Cathomen, T., Segal, D.J., 2011. Adding fingers to an engineered zinc finger nuclease can reduce activity. Biochemistry 50, 5033-5041.

Shirakawa, T., 2008. The current status of adenovirus-based cancer gene therapy. Mol Cells 25, 462-466.

Shrivastav, M., De Haro, L.P., Nickoloff, J.A., 2008. Regulation of DNA double-strand break repair pathway choice. Cell Res 18, 134-147.

Shukla, V.K., Doyon, Y., Miller, J.C., DeKelver, R.C., Moehle, E.A., Worden, S.E., Mitchell, J.C., Arnold, N.L., Gopalan, S., Meng, X., Choi, V.M., Rock, J.M., Wu, Y.Y., Katibah, G.E., Zhifang, G., McCaskill, D., Simpson, M.A., Blakeslee, B., Greenwalt, S.A., Butler, H.J., Hinkley, S.J., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., 2009. Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature 459, 437-441.

Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., Paques, F., 2011. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. Curr Gene Ther 11, 11-27.

Smih, F., Rouet, P., Romanienko, P.J., Jasin, M., 1995. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic Acids Res 23, 5012-5019.

Smith, J., Grizot, S., Arnould, S., Duclert, A., Epinat, J.C., Chames, P., Prieto, J., Redondo, P., Blanco, F.J., Bravo, J., Montoya, G., Paques, F., Duchateau, P., 2006. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. Nucleic Acids Res 34, e149.

Sokurenko, E.V., Tchesnokova, V., Yeung, A.T., Oleykowski, C.A., Trintchina, E., Hughes, K.T., Rashid, R.A., Brint, J.M., Moseley, S.L., Lory, S., 2001. Detection of simple mutations and polymorphisms in large genomic regions. Nucleic Acids Res 29, E111.

Sollu, C., Pars, K., Cornu, T.I., Thibodeau-Beganny, S., Maeder, M.L., Joung, J.K., Heilbronn, R., Cathomen, T., 2010. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. Nucleic Acids Res 38, 8269-8276.

Stephen, S.L., Montini, E., Sivanandam, V.G., Al-Dhalimy, M., Kestler, H.A., Finegold, M., Grompe, M., Kochanek, S., 2010. Chromosomal integration of adenoviral vector DNA in vivo. J Virol 84, 9987-9994.

Sugio, A., Yang, B., Zhu, T., White, F.F., 2007. Two type III effector genes of Xanthomonas oryzae pv. oryzae control the induction of the host genes OsTFIIAgamma1 and OsTFX1 during bacterial blight of rice. Proc Natl Acad Sci U S A 104, 10720-10725.

Sundquist, W.I., Krausslich, H.G., 2012. HIV-1 Assembly, Budding, and Maturation. Cold Spring Harb Perspect Med 2.

Swift, S., Lorens, J., Achacoso, P., Nolan, G.P., 2001. Rapid production of retroviruses for efficient gene delivery to mammalian cells using 293T cell-based systems. Curr Protoc Immunol Chapter 10, Unit 10 17C.

Szczepek, M., Brondani, V., Buchel, J., Serrano, L., Segal, D.J., Cathomen, T., 2007. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol 25, 786-793.

Tang, W., Ehrlich, I., Wolff, S.B., Michalski, A.M., Wolfl, S., Hasan, M.T., Luthi, A., Sprengel, R., 2009. Faithful expression of multiple proteins via 2A-peptide self-processing: a versatile and reliable method for manipulating brain circuits. J Neurosci 29, 8621-8629.

Thermes, V., Grabher, C., Ristoratore, F., Bourrat, F., Choulika, A., Wittbrodt, J., Joly, J.S., 2002. I-SceI meganuclease mediates highly efficient transgenesis in fish. Mech Dev 118, 91-98.

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L., Melton, D.W., 1989. Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell 56, 313-321.

Toietta, G., Mane, V.P., Norona, W.S., Finegold, M.J., Ng, P., McDonagh, A.F., Beaudet, A.L., Lee, B., 2005. Lifelong elimination of hyperbilirubinemia in the Gunn rat with a single injection of helper-dependent adenoviral vector. Proc Natl Acad Sci U S A 102, 3930-3935.

Tomanin, R., Scarpa, M., 2004. Why do we need new gene therapy viral vectors? Characteristics, limitations and future perspectives of viral vector transduction. Curr Gene Ther 4, 357-372.

Todaro, G. J., Green, H., 1963. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. J Cell Biol 17, 299-313.

Tordaro, G.J., Green, H., 1964. An assay for cellular transformation by SV40. Virology 23, 117-119.

Townsend, J.A., Wright, D.A., Winfrey, R.J., Fu, F., Maeder, M.L., Joung, J.K., Voytas, D.F., 2009. High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459, 442-445.

Trobridge, G.D., Kiem, H.P., 2010. Large animal models of hematopoietic stem cell gene therapy. Gene Ther 17,8, 939-48.

Tulman, E.R., Delhon, G.A., Ku, B.K., Rock, D.L., 2009. African swine fever virus. Curr Top Microbiol Immunol 328, 43-87.

Tzfira, T., White, C., 2005. Towards targeted mutagenesis and gene replacement in plants. Trends Biotechnol 23, 567-569.

Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., Holmes, M.C., 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435, 646-651.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., Gregory, P.D., 2010. Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11, 636-646.

van der Burg, M., Ijspeert, H., Verkaik, N.S., Turul, T., Wiegant, W.W., Morotomi-Yano, K., Mari, P.O., Tezcan, I., Chen, D.J., Zdzienicka, M.Z., van Dongen, J.J., van Gent, D.C., 2009a. A DNA-PKcs mutation in a radiosensitive T-B- SCID patient inhibits Artemis activation and nonhomologous end-joining. J Clin Invest 119, 91-98.

van der Burg, M., van Dongen, J.J., van Gent, D.C., 2009b. DNA-PKcs deficiency in human: long predicted, finally found. Curr Opin Allergy Clin Immunol 9, 503-509.

van der Burg, M., van Zelm, M.C., van Dongen, J.J., 2009c. Molecular diagnostics of primary immunodeficiencies: benefits and future challenges. Adv Exp Med Biol 634, 231-241.

van Hennik, P.B., Verstegen, M.M., Bierhuizen, M.F., Limon, A., Wognum, A.W., Cancelas, J.A., Barquinero, J., Ploemacher, R.E., Wagemaker, G., 1998. Highly efficient transduction of the green fluorescent protein gene in human umbilical cord blood stem cells capable of cobblestone formation in long-term cultures and multilineage engraftment of immunodeficient mice. Blood 92, 4013-4022.

van Meerloo, J., Kaspers, G.J., Cloos, J., 2011. Cell sensitivity assays: the MTT assay. Methods Mol Biol 731, 237-245.

Vaziri, H., Benchimol, S., 1998. Reconstitution of telomerase activity in normal human cells leads to elongation of telomeres and extended replicative life span. Curr Biol 8, 279-282.

Verma, I.M., Weitzman, M.D., 2005. Gene therapy: twenty-first century medicine. Annu Rev Biochem 74, 711-738.

Vigna, E., Naldini, L., 2000. Lentiviral vectors: excellent tools for experimental gene transfer and promising candidates for gene therapy. J Gene Med 2, 308-316.

Vile, R., 1992. The retroviral life cycle and the molecular construction of retrovirus vectors. Methods Mol Biol 8, 1-15.

Vink, C.A., Gaspar, H.B., Gabriel, R., Schmidt, M., McIvor, R.S., Thrasher, A.J., Qasim, W., 2009. Sleeping beauty transposition from nonintegrating lentivirus. Mol Ther 17, 1197-1204.

Vock, E.H., Lutz, W.K., Ilinskaya, O., Vamvakas, S., 1999. Discrimination between genotoxicity and cytotoxicity for the induction of DNA double-strand breaks in cells treated with aldehydes and diepoxides. Mutat Res 441, 85-93.

Vorburger, S.A., Hunt, K.K., 2002. Adenoviral gene therapy. Oncologist 7, 46-59.

Wagemaker, G., Hartong, S.C., Neelis, K.J., Egeland, T., Wognum, A.W., 1998. In vivo expansion of hemopoietic stem cells. Stem Cells 16 Suppl 1, 185-191.

Wakefield, J.K., Kang, S.M., Morrow, C.D., 1996. Construction of a type 1 human immunodeficiency virus that maintains a primer binding site complementary to tRNA(His). J Virol 70, 966-975.

Wanisch, K., Yanez-Munoz, R.J., 2009. Integration-deficient lentiviral vectors: a slow coming of age. Mol Ther 17, 1316-1332.

Warnock, J.N., Daigre, C., Al-Rubeai, M., 2011. Introduction to viral vectors. Methods Mol Biol 737, 1-25.

Weberpals, J. I.,Clark-Knowles, K. V.,Vanderhyden, B. C., 2008. Sporadic epithelial ovarian cancer: clinical relevance of BRCA1 inhibition in the DNA damage and repair pathway. J Clin Oncol, 26, 16, 3259-67.

Werner, M., Kraunus, J., Baum, C., Brocker, T., 2004. B-cell-specific transgene expression using a self-inactivating retroviral vector with human CD19 promoter and viral post-transcriptional regulatory element. Gene Ther 11, 992-1000.

Wiegant, W.W., Meyers, M., Verkaik, N.S., van der Burg, M., Darroudi, F., Romeijn, R., Bernatowska, E., Wolska-Kusnierz, B., Mikoluc, B., Jaspers, N.G., Vreeken, C., Ijspeert, H., Esveldt-van Lange, R.E., Friedl, A.A., de Villartay, J.P., Mullenders, L.H., van Dongen, J.J., van Gent, D.C., Pastink, A., Zdzienicka, M.Z., 2010. A novel radiosensitive SCID patient with a pronounced G(2)/M sensitivity. DNA Repair (Amst) 9, 365-373.

Williams, S., Mustoe, T., Mulcahy, T., Griffiths, M., Simpson, D., Antoniou, M., Irvine, A., Mountain, A., Crombie, R., 2005. CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnol 5, 17.

Wognum, A.W., Eaves, A.C., Thomas, T.E., 2003. Identification and isolation of hematopoietic stem cells. Arch Med Res 34, 461-475.

Wolff, J.A., Lederberg, J., 1994. An early history of gene transfer and therapy. Hum Gene Ther 5, 469-480.

Wong, G.K., Chiu, A.T., 2011. Gene therapy, gene targeting and induced pluripotent stem cells: applications in monogenic disease treatment. Biotechnol Adv 29, 1-10.

Wong, H.H., Jiang, G., Gangeswaran, R., Wang, P., Wang, J., Yuan, M., Wang, H., Bhakta, V., Muller, H., Lemoine, N.R., Wang, Y., 2012. Modification of the early gene enhancer-promoter improves the oncolytic potency of adenovirus 11. Mol Ther 20, 306-316.

Yamashita, M., Emerman, M., 2006. Retroviral infection of non-dividing cells: old and new perspectives. Virology 344, 88-93.

Yanez-Munoz, R.J., Balaggan, K.S., MacNeil, A., Howe, S.J., Schmidt, M., Smith, A.J., Buch, P., MacLaren, R.E., Anderson, P.N., Barker, S.E., Duran, Y., Bartholomae, C., von Kalle, C., Heckenlively, J.R., Kinnon, C., Ali, R.R., Thrasher, A.J., 2006. Effective gene therapy with nonintegrating lentiviral vectors. Nat Med 12, 348-353.

Yanez, R.J., Porter, A.C., 1998. Therapeutic gene targeting. Gene Ther 5, 149-159.

Yang, B., Wen, X., Kodali, N.S., Oleykowski, C.A., Miller, C.G., Kulinski, J., Besack, D., Yeung, J.A., Kowalski, D., Yeung, A.T., 2000. Purification, cloning, and characterization of the CEL I nuclease. Biochemistry 39, 3533-3541.

Yoder, K.E., Fishel, R., 2008. Real-time quantitative PCR and fast QPCR have similar sensitivity and accuracy with HIV cDNA late reverse transcripts and 2-LTR circles. J Virol Methods 153, 253-256.

Young, J.J., Cherone, J.M., Doyon, Y., Ankoudinova, I., Faraji, F.M., Lee, A.H., Ngo, C., Guschin, D.Y., Paschon, D.E., Miller, J.C., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Harland, R.M., Zeitler, B., 2011. Efficient targeted gene disruption in the soma and germ line of the frog Xenopus tropicalis using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A 108, 7052-7057.

Yucel, R., Kosan, C., Heyd, F., Moroy, T., 2004. Gfi1:green fluorescent protein knock-in mutant reveals differential expression and autoregulation of the growth factor independence 1 (Gfi1) gene during lymphocyte development. J Biol Chem 279, 40906-40917.

Yun, M.H., Hiom, K., 2009. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. Nature 459, 460-463.

Zanta-Boussif, M.A., Charrier, S., Brice-Ouzet, A., Martin, S., Opolon, P., Thrasher, A.J., Hope, T.J., Galy, A., 2009. Validation of a mutated PRE sequence allowing high and sustained transgene expression while abrogating WHV-X protein synthesis: application to the gene therapy of WAS. Gene Ther 16, 605-619.

Zhang, F., Frost, A.R., Blundell, M.P., Bales, O., Antoniou, M.N., Thrasher, A.J., 2010a. A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors. Mol Ther 18, 1640-1649.

Zhang, F., Maeder, M.L., Unger-Wallace, E., Hoshaw, J.P., Reyon, D., Christian, M., Li, X., Pierick, C.J., Dobbs, D., Peterson, T., Joung, J.K., Voytas, D.F., 2010b. High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. Proc Natl Acad Sci U S A.

Zhao, L., Bonocora, R.P., Shub, D.A., Stoddard, B.L., 2007. The restriction fold turns to the dark side: a bacterial homing endonuclease with a PD-(D/E)-XK motif. EMBO J 26, 2432-2442.

Zibara, K., Hamdan, R., Dib, L., Sindet-Pedersen, S., Kharfan-Dabaja, M., Bazarbachi, A., El-Sabban, M., 2012. Acellular bone marrow extracts significantly enhance engraftment levels of human hematopoietic stem cells in mouse xeno-transplantation models. PLoS One 7, e40140.

Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K., Chen, G., Ye, Z., Park, I.H., Daley, G.Q., Porteus, M.H., Joung, J.K., Cheng, L., 2009. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell 5, 97-110.

Zufferey, R., Donello, J.E., Trono, D., Hope, T.J., 1999. Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. J Virol 73, 2886-2892.

Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., Trono, D., 1998. Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J Virol 72, 9873-9880.

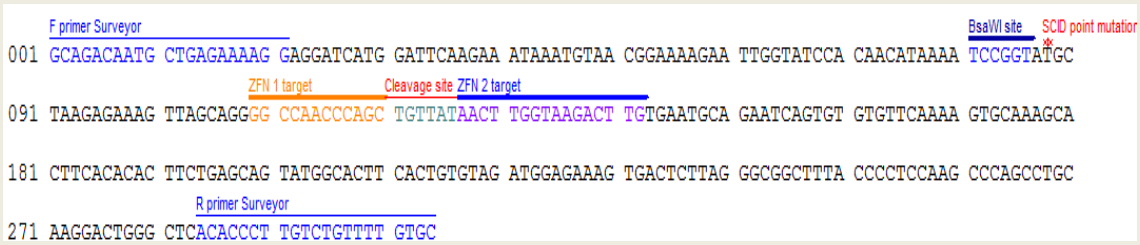
Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L., Trono, D., 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol 15, 871-875.

Zykovich, A., Korf, I., Segal, D.J., 2009. Bind-n-Seq: high-throughput analysis of in vitro protein-DNA interactions using massively parallel sequencing. Nucleic Acids Res 37, e151.

# Appendix I

Box 1

The sequence of Surveyor PCR product of (of 304 bp). Sequence F and R Surveyor primers are shown at the 5' and 3' end of the amplicon sequence. The ZFN 1 target is the recognition site for **17834-*FOK* EL ZFN** domain; whereas ZFN 2 target is the recognition site of **17373-*FOK* KK ZFN** domain and in between there is the cleavage site where the DSB should take place. The diagnostic *Bsa*WI restriction site is located upstream of the *scid* point mutation which is in turn upstream of the ZFN full recognition site.



Box 2

The sequence of *Prkdc-Bsa*WI corrective template (of 1626 bp) amplified by PCR. Sequence F and R *Bsa*WI assay primers are shown at the 5' and 3' end of the amplicon sequence. This template also includes the sequence and features of Surveyor PCR product of (304bp) shown in Box 1 above.

F primer BsaWI assay

0001 GACCCAATGT TTAGTTTTTA TGAGTGTTTT GCCTGCAGGT ATGTAAGTGC ATCATATGCA TGCAGTGCCC ACAAAGTCCA GAAGAGGGAG

0091 CACCTGAATT AGAATTACAG AAGATGTAAC CAATGTTTCT GGGACCTGAA TGCTGGTTGG TCCTCTGTAA GAACATCAAG TGCTCTTAAT

0181 CCTTGAACCA TCTCTCCAGC CAACAACAAC CTTTGATTTA AAAATCATTC CATGTGAATT TTTCCTTAAT TTCCTTATCT AAGAACAGAT

0271 AATGCTGTTT GACTCAATTA TCCAGACTAT CCCCGAAAGC ATTAAGCTCA TCTTTTTATG TGGACTTCTT CTCACCTCTC TGCATCGCAT

0361 GATCATGGTG ATATGTGCCA TGTCAGTGTC TGACTAGAAA GCTAGAGAGC TGTTCCAGTT ATAGATCTTT GTTTTAGGGT CATTACTTGG

F primer Surveyor

0451 TTTAATGTTT TTTAATGTAA TTTGTATATG CTATTATAAT AAGTAGAAAA AAATGTGTTT TTTCCCTTAG AGTTTTGAGC AGACAATGCT

(F primer Surveyor) BsaWI site SCID point mutation

0541 GAGAAAAGGA GGATCATGGA TTCAAGAAAT AAATGTAACG GAAAAGAATT GGTATCCACA ACATAAAATC CGGTATGCTA AGAGAAAGTT

ZFN recognition site

0631 AGCAGGGGCC AACCCAGCTG TTATAACTTG GTAAGACTTG TGAATGCAGA ATCAGTGTGT GTTCAAAAGT GCAAAGCACT TCACACACTT

0721 CTGAGCAGTA TGGCACTTCA CTGTGTAGAT GGAGAAAGTG ACTCTTAGGG CGGCTTTACC CCTCCAAGCC CAGCCTGCAA GGACTGGGCT

R primer Surveyor

0811 CACACCCTTG TCTGTTTTGT GCTTCTGTAC AGTCAGCACC AATAGGAGAA CTAAGGTCAT GCCATAATAT TGTTATGCAT ATGAGGCTTG

0901 GAACTCATAT TTGCCTTTGT ACTTTATAGG GCATTTGGCC CACTTGATCC TGATGAACAC TAGTGTCAGG CATTTGACTA CTATCTCATA

0991 GTGAAAACAC AACGTGGGCT TCTCAGGTAG GAACCTGAGC TGCATTGACC TTTAGGTTTG CTTGAAGTTT TTAAAAATTG AAATCTAAAG

1081 TAAACAGAAA TGTGGAGGTA TTATTCCTAG GAAGTAAAAA CCTTAGACTT CAGTTCTGAG CAGTAGTAAG ATGTACTCTG TAATTGACTT

1171 GATTTTTCAC AATAGTGTCA GATCTATACA GTGACTGGTG CTAGTAAGAG AATGAGCTGC AGAAATCAAT TCAAAGTCTA TCAGACTTAA

1261 ATGTAAAACA TTAAGTTGAT CTTCATGATG GCACTTGGCT GTGACATAAC ATTGCTAGGC ATGACCCCAG AAGCACGAGG AAGACCAAGA

1351 CACACAAGCT GGACTTGAGA ATTCCCAGCT GTGCTTTACA GGGTGTAGGC AAGAACAGAG GAGGACTGGG ACAGTAAGTG AGGGTTCCTG

1441 AGCTAATCTA AAAGCACAAG GCCTGGCTTC AATTCCCAGA GCTAGATTTT GTTTTAAAAT TAGGGTTGCA GAAATGACTT AGCAATAAGA

R primer BsaWI assay

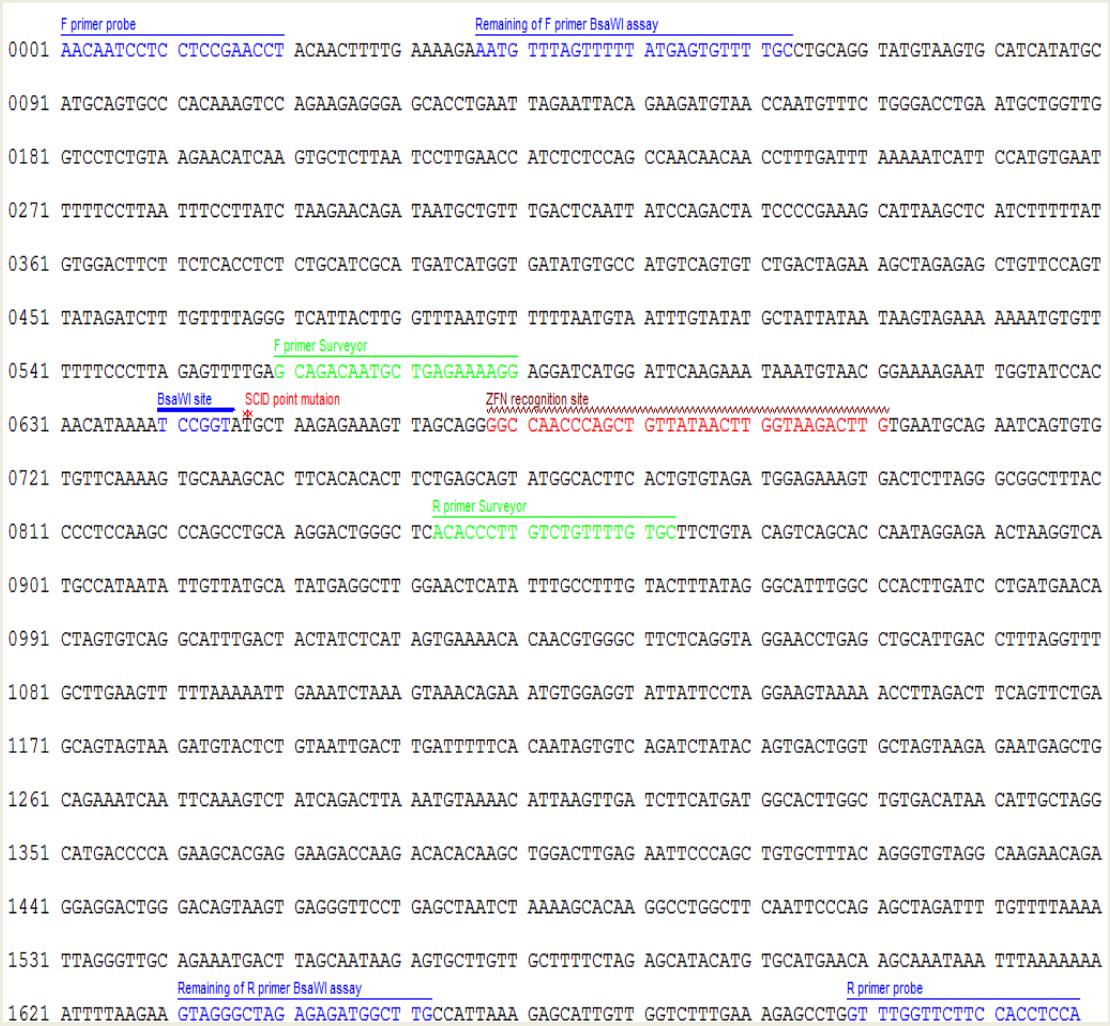
1531 GTGCTTGTTG CTTTTCTAGA GCATACATGT GCATGAACAA GCAAATAAAT TTAAAAAAAA TTTTAAGAAG TAGGGCTAGA GAGATGGCTT

(R primer BsaWI assay)

1621 GGGGTC

Box 3

The sequence of *Prkdc* probe (1708bp) amplified by PCR. Sequence F and R probe primers are shown at the 5' and 3' end of the amplicon sequence. The template also includes the sequence and features of *Prkdc* corrective template (1626bp) shown in Box 2 above.



Box 4

The sequence of gene targeting product (of 1335bp) resulted from co-transfection with ZFNs and p*Prkdc*Hind*Pml*I*neo*F constructs created by PCR. Sequence F and R probe primers are shown at the 5' and 3' end of the amplicon sequence. The template also included the recognition sites for restriction enzymes *Bgl*II, *Xho*I and *Hind*III.

F primer of gene repair

0001 TCGCCTTCTT GACGAGTTCT TCTGAGGGGA TCAATTCTCT AGAGCTCGCT GATCAGCCTC GACTGTGCCT TCTAGTTGCC

0081 AGCCATCTGT TGTTTGCCCC TCCCCCGTGC CTTCCTTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT

0161 GAGGAAATTG CATCGCATTG TCTGAGTAGG TGTCATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA AGGGGGAGGA

0241 TTGGGAAGAC AATAGCAGGC GATAACTTCG TATAGCATAC ATTATACGAA GTTATCAATT CGATGTGAAG GCTCACAACC

0321 ATCTGTACAG CTACAGTGTA CTCATATACA TAAATAAATC TTTTTTTAAA AAAAAGGGCA GTCCACATAT TGACAGAAAT

0401 CTTCAAGTCA TATTATTTGA GAATAGATTC ATGCCTAAAA TAAGTTTATG AGTTTTTTAA AGGCACGGTG TGGGGCCAAG

0481 GGGTGATAGA GAGACAGCTC CGTAGTTAAG AGCATTGCTC TTTCAGAGGG CTAGAGTTCT AGTTCTCATT TTACAAACCT

BglII

0561 TATAACCACT TGTAATTCCA ACTCCTAAAG ATCTGATGTC CTCTTCTGGA CCCTGGAGGC ACTCACAAGC ACAGACACAA

0641 ATCTCTAAAA TAATTTTCCT TAAAACCCAG TGTGGTGATA CAAACCTGTA ATTTCAGCAG ATGGGAGGTA GACGCCATCC

0721 TTGACTATAT ATACACTATA TATATATATT TTTTTAAAAC ATGATCAAAG TAATCAAACA TTTTTAAAAC ATGATCAAAG

0801 TAATCAAACA TTTTTCTTTT TGGTTTTTAG AGACAGGGTT TCTCTGTATA GCCCTGGCTG TCCTGGAACT CACTCTGTAG

0881 ACCAGGCTGG CCTCGAACTC AGAAATCCGC CTGCCTCTGC CTCCCGAGTG CTGGGATTAA AGGCGTGTGC CACCACGCCC

XhoI

0961 GACTCAAACA TTTTTCTAAA GAAAAATGTA CAAATGGTAC AAATGGCCAC TAACTCGAGG GAGACGTACT CAGCATTAGT

1041 AGTAGTTTAG CAAATGCAAT GACAATCCGC ATCTTGTGTG TTATGACAGC TTTCAGATGA GATGAAGTGT GAGCAGGGGT

1121 GTGGAGAGCT GCATGGATGG TGGGCATGGA AATTTAAATA GAGTTACAAT TGTAGCCAGC ATGCCCACTC CTGAGTGGCT

Hind III

1201 ACCCTAGAAA AATGAAAGCT TATGTTTGAA CAAAAATTTC CAAGCAAATG TTAATAATAA TGTTATTATT CATGCTTTAT

R primer of gene repair

1281 TTTTCTACCA ACCCAAACTT TGAATTCACA GTCATGAGTG ACATGAGGGG GAAAA

Box 5

Table of home made / obtained plasmids that were used throughout current study.

| **Plasmid** | **Made by/ obtained from** |
|---|---|
| pCCLsc ISce-IT | Home made by Rafael Yáñez |
| pCCLsc_S_W | Home made by Rafael Yáñez |
| pHR'sc-SEW | Obtained from Prof Adrian Thrasher (UCL) |
| pPrkdcHindBsaWI | Home made by Rafael Yáñez |
| pRRLsc_C_W | Home made by Rafael Yáñez |
| pRRLsc-CEW | Obtained from Prof. Luigi Naldini (Milano, Italy) |
| pVAX-3FN-17373-FokKK | Made by Sangamo Biosciences Inc. (California-USA) |
| pVAX- N2A-3FN-17834-FokEL2 | Made by Sangamo Biosciences Inc. (California-USA) |

Box 6

Table of collaborators and their detailed contributions to the currnt study.

| **Collaborator** | **Contribution** |
|---|---|
| Dr. Céline Rocca (RHUL) | Doing together some cloning, *Bsa*WI assay, DNA-PK activity assay and melphalan enrichment |
| Dr. Manfred Schmidt (NCTD, Germany) | Deep sequencing of potential samples |
| Dr. Maria Alonso-Ferrero (ICH-UCL) | Setting up of multicolour flow cytometric analysis. |
| Dr. Michael Blundell (ICH-UCL) | Injection and bleeding of transplanted animals. |
| Dr. Ramon Alemany (IDIBELL, Spain) | Preparation of adenoviral vectors. |
| Miss. Sara Oliván (RHUL) | Preparation of some lentivectors (GFP IDLV with different promoters). |
| Sangamo Biosciences (CA, USA) | Design and preparation of ZFNs. |